

How to neutralize acid rain

The clamour for action — almost any action — will stampede governments into foolishness. The problem is more complicated than it seems, while the legal framework for control is neglected.

MUCH of Europe has spent much of this summer worrying about rain. In some places (such as western Britain) there has been too little of it. Elsewhere, people seem to have taken fright that what little rain there has been is dangerously acidic. Acid rain is well on the way to being in the 1980s a focus for general discontent with technology as powerful, and ultimately as cathartic, as the war waged against pesticides (and their manufacturers) in the 1960s. The danger, now as then, is that people will compel their governments to jump to hasty and even exaggerated conclusions about the remedies that should be adopted.

None of this is intended to imply that there is no problem to be dealt with. On the contrary, there is ample proof that the ingredients of a problem of some kind do exist. Burning fossil fuels is a well-known way of putting sulphur dioxide into the atmosphere and there are well-known mechanisms by which, in the presence of sunlight and oxygen, it can be converted into sulphuric acid. The annual production of sulphur dioxide is now well over 100 million tonnes for the world as a whole. Moreover, since the careful study carried out by the Organization for Economic Cooperation and Development (OECD) a decade ago (published in 1977), it has been clear that what Scandinavians and Canadians had previously only suspected is indeed true — that atmospheric pollution does not recognize national frontiers. What seems to have dramatized the issue of acid rain in Europe is the recognition, in West Germany and Switzerland in particular, that trees in previously healthy forests are dying off without obvious explanation. Especially where "Greens" have political influence, it is inevitable that public attention (and a good many headlines) should be concerned with acid rain. The clamour for a "solution" is deafening but the most urgent need is to understand what the problem is.

Lakes

It is now forty years since Dr S. Oden at the University of Uppsala drew attention to the increasing acidity of Swedish lakes, and to the decrease of fish populations in some of them. Oden's conclusion was that the phenomena are related as cause and effect. If acidity does not kill fish directly, it may do so by leaching toxic metals from the bedrock. Since then, the acidity of many lakes has increased steadily in the course of time. The difficulty, even in Scandinavia, is that lakes are not equally susceptible to whatever may be happening to them. In a largely igneous terrain, natural acid-base buffering is often ineffectual and frequently much less than it would be in, say, a limestone area. Nobody will be surprised if the acidity of some Scandinavian lakes is increasing for precisely the reasons at which Oden guessed, but the circumstance that some are comparatively robust needs an explanation not so far forthcoming. Since most of the rainwater reaching lakes and water-courses is first filtered through vegetation and soil, it would also help if there were some understanding of why the natural filter fails to hold back more of the acidity. F.T. Last (*Forstwissenschaftliches Centralblatt* 103, 24; 1984) is not the only one to have raised the possibility that the practices of foresters, and especially their fondness for fast-growing species such as spruce, may be partly responsible.

What does, however, seem clear is that the phenomena observed in Central Europe are differently caused. The guess that forest trees are being damaged by ozone in the atmosphere, itself

formed from nitrogen oxides, may not be very wide of the mark. In these circumstances, tighter regulation of sulphur dioxide emission may contribute very little to a lasting solution. The urgent need is to understand what is happening to these forests. In a helpful analogy with Koch's consideration of the mechanisms of infection in animal disease, Last points out that in the study of the relationship between atmospheric pollution and the health of forest trees, there have so far been very few attempts to demonstrate that presumptive agents of damage do indeed function in the ways supposed.

But is not the cry for more pointed research the standard recipe for delay? This fear has some substance. Governments and electricity utilities (the chief offenders on sulphur dioxide) have over the years been too quickly dismissive of the accumulating evidence. The accusation that they allow vested interest in cheap electricity production to get the better of their judgement is probably, however, misplaced. Unlike most other businesses, electricity utilities can pass on to their customers such extra costs as they may be required to shoulder — that they should install scrubbers at their plants, for example. In the present climate, they have no cause single-handedly to stand out against the acid rain campaign. A better understanding will be in everybody's interest.

Research

In Britain, the nationalized Central Electricity Generating Board and the National Coal Board have contributed to the £5 million cost of a research programme mounted by the national academies of Britain, Norway and Sweden. (The worst outcome, a conclusive proof that British sulphur dioxide is causing serious damage in Scandinavia, would probably be welcome to the electricity utility, eager as it is to build more nuclear generating capacity.) Elsewhere in Europe, most governments are supporting newly devised research projects, while the European Science Foundation is planning to mount a study of the effects of atmospheric pollution on Central European forests. Much the same is happening in North America, although it remains a crying shame that the project for a joint research project between the United States and Canada has been stymied by the US Administration's fears that it would be bound by the implications of the outcome. What tends to be forgotten, in the clamour for quick action on sulphur emissions, is that the social cost in the polluting countries could be very large. An extra twenty per cent on the cost of electricity would be no joke.

It is also too easily forgotten that the technical problems of dealing with acid rain are the least of those that lie ahead. In Europe, Scandinavian governments have been asking since the early 1970s that something should be done to control the emission, from elsewhere in Europe, of sulphur and nitrogen oxides, but without much success. The most tangible success so far is the European Commission's directive (in principle binding on all member states), promulgated last year, that sulphur dioxide in ambient air should in future not exceed 120 micrograms per cubic metre. The snag, for Scandinavia, is that the standard may be met either by reducing the output of sulphur dioxide or by building taller chimney stacks to ensure that pollution is more widely dispersed. In any case, governments now offending have until 1993 to comply. In North America, even less has been done to meet Canadian demands for a limit to the output of sulphur



dioxide from the northwestern states. The reasons are easily understood, if not conspicuously creditable: a government that agrees to limit its production of sulphur dioxide for fear of the damage that might be done in neighbouring states is voluntarily abandoning some element of sovereignty, usually cherished as indivisible.

This is why the most serious problem thrown up by anxiety over acid rain is not so much the technical question of precisely what damage is done by atmospheric pollution but that of the legal framework in which the damage may be accommodated. Precedents are few, most amply represented by the fisheries agreements under the terms of which governments undertake to restrain their nationals from behaving in a way that harms an international resource. But there is already in place the framework of an international convention on atmospheric pollution beneath the improbable umbrella of the United Nations Economic Commission for Europe (which, again improbably, includes both Eastern European states and the Soviet Union and also the United States and Canada). This legal instrument owes its existence to a declaration by Mr Leonid Brezhnev in Stockholm during the negotiation of the Helsinki agreements on European security, but its provisions are for the time being properly flexible. The signatories do little more than acknowledge that acid rain may occasion transnational problems which, if discovered, may require restraint or the payment of compensation. Fortunately, however, the convention includes sensible provisions for taking account of new understanding. The first need is to find out for sure what needs to be done, the second to pay more attention to the potential of this unique agreement, potentially an effective international regulatory instrument. □

Chinese example

The rapid fall in China's birth-rate is an eye-opener to others.

THE United Nations conference in Mexico City on world population seems to have passed off less acrimoniously than might have been the case (see *Nature* 9 August, p.439). The delegation of the United States made its advertised stand on abortion, declining to provide direct subventions for organizations advocating abortion as an instrument of population control, but not so fiercely as to prevent the conference from adopting an agreed statement, the highest common factor of success at these occasions. To the extent that the form of words agreed also acknowledges the intimate inverse relationship between economic and social development on the one hand and population growth on the other, Mexico City may yet mark the emergence of a more level-headed regard for the problem of world population. The conference will also have served to dramatize for governments of all kinds the most striking lesson to have been learned in the past decade — that the demographic transition (from a state of balance at high birth and death rates to one where both are low) need not be a slow process. The recent demographic history of the People's Republic of China has people transfixed.

By any standards, the changes that have been brought about in the demographic pattern in China are staggering, with the reduction of the death rate by far the most conspicuous. Chinese sources say that the crude death rate fell from 13.2 a thousand in 1952 to 6.2 a thousand in 1979. Even if the second figure must be taken with a pinch of salt, there seems no doubt that the rate has fallen to a half in just over a quarter of a century and, perhaps more important, that the disparity between the urban and rural populations has been substantially removed. A larger fraction of newborn females survive to childbearing age, while the expectation of life at birth of both sexes is increased, to close on three score years and ten. While it remains a puzzle for the outside world that the Chinese Government has been able, on paper at least, to promulgate (in 1979) the rule that each family must in future have no more than one child, it seems unlikely that this draconian exercise could have been mounted at all if there had not

been such a dramatic improvement in life expectation.

Meanwhile — and this is what the rule is for — the birth-rate is declining. The usual conical shape of the age distribution of a population is pathetically indented in the case of China on two occasions, forty and twenty years ago. The first disturbance corresponds with the years of the Long March and the communist revolution, the second to the self-inflicted injury whimsically called the cultural revolution. Remarkably, however, the birth-rate began falling rapidly (from a peak of close on 45 per 1,000) once the cultural revolution ended. Between 1969 and 1979, it halved. But will the Government of China be able to hold the line it has drawn in the social sand, and for how long? Most probably the economic penalties for having more than one child will seem less onerous to some sections of the population as the country becomes more prosperous. □

Freedom to travel

The European Physical Society should have resisted Prague's restrictions.

THE trouble that has blown up about the European Physical Society's meeting planned for Prague later this month (see p.617) raises general questions that deserve wide attention. The circumstances are that Dr F. Janouch, once a Czechoslovak citizen and now a citizen of Sweden where he lives, has been denied re-entry to Czechoslovakia so as to attend a meeting of the council to which he was elected, and of which the Czechoslovak Academy of Science is also a member. Professor John Ziman, another member, has also as a protest rightly declined to attend. The society says it is too late to rearrange the meeting.

The question that needs to be considered by the European Physical Society and others is the proper definition of the circumstances in which they will go to the stake in defence of a member's legitimate right to share in an agreement to travel to places where freedom is not guaranteed. Traditionally and rightly, organizations such as the International Council of Scientific Unions have fought to establish that meetings labelled as international should be open to all *bona fide* scientists. Broadly speaking, these efforts have succeeded. Particular difficulties often crop up, but most international meetings pass off without difficulty.

How far should societies go in their defence of members' rights of participation? Where a society claims the right of one of its members to travel to a foreign venue, it would be disingenuous also to claim that the person concerned should be immune from prosecution for some criminal offence of which he had been lawfully indicted, perhaps on some earlier visit. Difficulties arise, however, when a host government such as that of Czechoslovakia regards an unwelcome prospective visitor as an enemy of the state on grounds that fall short of the penal code, which appears to be Janouch's case. A further difficulty is that it may seem unreasonable that people in his position, separately denied the right to return to their home ground, should be able to return under the aegis of an international organization. That is irrational — and the European Physical Society should have dug in its heels at Prague's refusal. By the same test, however, the society should have been willing to undertake, on its own behalf and Janouch's, that he would not take the opportunity to rub salt into Prague's wounded pride.

Whether the dispute could have been settled on such a basis — and whether the terms of such an agreement would have been acceptable to Janouch — will not now be known. What is clear, however, is that cases like these are unhappily likely to be more and not less common in the years ahead. The International Council of Scientific Unions should give some thought to the question at its meeting at Ottawa next month. It should also worry a little about a less explicit and thus potentially more insidious threat to the freedom of communication — devices such as that by which the Government of India makes travel to India difficult for Israelis by placing obstacles in the way of visa applications. To adapt a phrase from common-market jargon, this is a non-visa restraint on travel and should equally be resisted. □

Recombinant DNA

Prospect of new US regulation

Surprise welcome for EPA

Washington

THE White House and the biotechnology industry appear to be dropping earlier objections to the Environmental Protection Agency (EPA)'s plan for regulating genetically-engineering microorganisms. A new draft proposal, completed by the agency in June and scheduled for publication this autumn, has received generally high marks for setting out a logical approach for filling the gaps in the authority of the Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health (NIH).

The largest of these gaps is in RAC's authority — or formal lack thereof — over commercial activities; NIH's rules are binding only upon recipients of federal research grants. Although no companies have yet been known to circumvent the RAC rules, that may change soon, especially in the wake of the temporary restraining order granted to anti-genetic-engineering activist Jeremy Rifkin that prevents RAC from considering proposals for the field-testing of recombinant organisms. The order does not apply to commercial proposals and, indeed, RAC endorsed two such at its June meeting, although NIH Director James Wyngaarden is not expected to approve the recommendation. Wyngaarden is said to take seriously Rifkin's objection that to approve commercial field tests while federally-sponsored field tests are blocked would create a double standard.

The result could be a delay of several years for companies that continue voluntarily to comply with RAC rules. No trial date has been set for Rifkin's claim that RAC must file an environmental impact statement before taking up field test proposals; if a decision is made in Rifkin's favour, it would take at least another year for RAC to prepare that document.

If for no other reason than the lengthy delays in prospect, industry is looking favourably on EPA's plans to assert jurisdiction. The new draft proposal, copies of which have been widely circulated in recent weeks even while officially secret, seems also to have reassured the industry and the White House that EPA has made a serious effort to understand the scientific issues involved; the same could not be said for an earlier draft prepared in March.

Industry will have little reason to complain of overburdensome regulation, either. The principal proposals would require that EPA be notified in advance of planned releases of genetically modified microorganisms into the environment.

Organisms intended to act as pesticides, whether natural or modified, already require EPA approval of safety and effectiveness before they can be marketed. Present rules, however, allow a blanket exemption for field testing on plots smaller than 10 acres. The agency is proposing to eliminate that exemption for genetically modified microorganisms, but wants to stop short of requiring a formal application (for an "Experimental Use Permit") for every such test. Companies would simply notify EPA of the planned test and provide basic data; if the agency does not object within 90 days, the tests could go ahead.

For other organisms, EPA is planning to regulate environmental releases under the Toxic Substances Control Act (TSCA), by defining modified DNA as a "new chemical substance". "Newness" would be determined operationally by taking account of the process used to produce it, thus avoiding the difficulties which arise because DNA recombination occurs in nature. TSCA basically requires manufacturers to notify the EPA before manufacturing any new chemical. The agency can hold up manufacturing if it believes additional safety data are needed and can, ultimately, seek a ban or restriction on the chemical.

As under the pesticide regulations, an existing exemption for research and development would be modified to restrict field testing of modified organisms. Recombinant DNA research confined to the laboratory would remain within the purview of RAC.

The TSCA rules would not apply to plants or animals, or to organisms used to make foods, food additives or drugs. They would, however, apply to engineered organisms used in the production of other chemicals — even if they were confined to a laboratory or manufacturing plant. The industry is expected to protest at this feature of its proposed rules, which would probably have an immediate effect on the plans of the biotechnology companies.

As things are, the companies have on their books only a few projects entailing the release into the environment of engineered organisms — chiefly bacteria for dealing with pollution and for symbiotic nitrogen fixation.

Although the agency says that it is not now planning to regulate naturally-occurring organisms in environments other than those in which they naturally occur, it notes that it could claim authority to do so under the "significant new use" provisions of TSCA.

Stephen Budiansky

Space experiments

Hearts in mouth at launch delay

EUROPEAN space scientists were relieved, earlier this week, that their tripartite experiment (with the United States) was launched last week from Cape Canaveral, one week late. In a field in which improvisation is frowned on, the circumstance that last-minute work with a vacuum cleaner was necessary to get the three satellites into orbit it counted a desperate measure.

The objective of the experiment, which consists of three satellites called the Active Magnetospheric Particle Tracer Explorers (AMPTE), is to follow the behaviour of ions in the Earth's magnetosphere. To this end, quantities of lithium and barium will be released at three pre-planned intervals during the next twelve months. Part of the anxiety about the prospect that last week's launch might be delayed stemmed from the knowledge that the first release is planned to coincide with the coming equinox.

The experiment's three satellites were provided by West Germany (to release lithium), the United Kingdom (to follow released material at close quarters) and the United States (to carry out observations at a distance). The orbits successfully attained last week are highly eccentric, ranging between 550 km and nine Earth radii (175,000 km). The apogee of the British and West German satellites is to be increased to 18.7 Earth radii before the first release (of lithium), allowing material to be injected into the solar wind well outside the magnetopause.

The US satellite is equipped for the recognition of both lithium and barium ions but also for the monitoring of the normal constituents of the ions trapped within the magnetosphere. The first phase of the experiment should throw some light on the transmission of ions from the solar wind through the magnetopause.

By the later stages of the experiment, the apogee of the two more distant satellites will by precession lie between the magnetopause and the Earth's bow shock of the solar wind beyond it (in December this year) and in the downwind tail of the magnetosphere (at next year's spring equinox). There is particular interest in the last of these releases (of a mixture of lithium and barium), when it should be possible to learn something of the behaviour of relatively dense plasma within the trapped plasma of the magnetosphere.

The cause of the delay that could have postponed the experiment for six months remains obscure. After the failure of a West German tracking computer, the space housing the three satellites (stacked on top of each other) was found to have been contaminated with material from the lining of an umbilical connection used for flushing the satellite compartment with nitrogen. □

Japanese research

Ministry's ambitions still higher

EFFORTS to increase Japan's research and development expenditure to a staggering 3 per cent of gross national product (GNP), compared with the present 2.44 per cent (a couple of tenths of a per cent more than most Western nations) are once again being called for by the Ministry of International Trade and Industry (MITI).

The "three-per-cent-of-GNP" goal was first proposed in 1980 and its re-emergence at a time when the government is calling for more severe budget restraint than ever before is a symptom of MITI's current "high technology" fever: the belief that the holy trinity of "biotechnology, electronics and new materials" will guarantee Japan's prosperity into the twenty-first century. According to MITI estimates, electronics and information processing equipment will have a market of around 24 million million yen (£74,000 million) in 1990 and new materials and biotechnology Y10 million million and Y4.7 million million respectively in AD 2000.

To attain the three per cent level would require a massive change in the way research and development in Japan is supported. At present, the government provides only 27 per cent of the nation's research funds, while the rest of the money comes from (and is spent by) industry, in stark contrast with the United States and the United Kingdom where a little over 50 per cent of support is put up by government. Even after allowing for the very much higher level of government-sponsored defence research in the United States, the Japanese Government still provides relatively less support for research (25 per cent against the US 33 per cent). The

three per cent level would, however, change all that, requiring the government to take a 40 per cent share in the nation's research budget.

Such an increase seems scarcely possible, but MITI believes that Japanese industry cannot carry out the large-scale basic research needed in high technology without substantial government support. And alongside a simple direct increase in funding of the government's own research laboratories, MITI is proposing a series of moves to stimulate "high-tech" industrial research — and to link more tightly government and industrial research.

To bolster public support for the expansion of high technology research, MITI has made much of recent reports of the increasing "American threat": news that the industrial plant and equipment in the United States is now on average newer than in Japan and that US capital investment is increasing while that of Japan is decreasing.

MITI's wild enthusiasm for high technology is at least partly motivated by a desire to restore the prestige and influence it enjoyed in the high growth era of the 1960s. The industries it fostered then and which were the locomotives of the Japanese economic miracle — steel, petrochemicals and paper — are now in decline. Despite recent MITI efforts to enter into the rapidly expanding telecommunications field, the Ministry of Posts and Telecommunications remains in control. MITI is now looking to high technology to save its own prestige as well as to guarantee Japan's future prosperity.

MITI's war cry as it drafts next year's budget request is bound to sharpen ten-

sions with the Ministry of Education, Science and Culture (MESC), the chief source of support for the universities and their "basic science". MITI officials are openly dismissive of MESC for failing to get the universities to work effectively with industry or to move quickly enough as research priorities change. But to those in the universities it seems that MITI policies reduce their role to that of supplying trained personnel to industrial laboratories — and even that is not an easy job.

In biotechnology, for example, there seems little recognition of the need to maintain a large pool of researchers doing truly fundamental research on the many diverse specializations that will need to be brought to bear for success. Although money for molecular biology research is available, there is a considerable shortage of staff.

Indeed, it is doubtful whether Japanese biotechnology companies could have reached their present stage of development without the return of many scientists who were working in laboratories in the United States and Europe. When it comes to the science budget, though, the power and influence of MITI (and the Science and Technology Agency) outweighs that of MESC, and there is little possibility, even if major increases in research spending are made, that the universities will benefit much from them.

Alun Anderson

AIDS

London scare ends

BLOOD tests of London laboratory technicians thought to have been exposed to acquired immune deficiency syndrome (AIDS) virus when a vacuum pump leaked have proved negative. The scare occurred at the Middlesex Hospital early in June when the retrovirus HTLV-III was being coated on plastic for antibody tests.

It was claimed by the trade union ASTMS (Association of Scientific, Technical and Managerial Staffs) that a vacuum pump used in this process was deficient. Professor John Patterson, head of the hospital's virology department, has denied this, saying that the pump was outside the Class 1 safety cabinet in which the plating experiment was being carried out. He said that it was not possible to guarantee that an aerosol had not been discharged into the air, but that it was very unlikely. There were three traps in the pump line, the last one a dry trap, which had remained dry.

Patterson says that it was right that ASTMS should be concerned, but that their fears had been dispelled. Mr Pike, safety representative of ASTMS at the Middlesex Hospital reported that a meeting between ASTMS and hospital representatives agreed that safety procedures would be reviewed and the pathology laboratory involved upgraded to full Category 3 status.

Marcus Chown

MITI's proposals

THE Ministry of International Trade and Industry (MITI) is basing this year's budget proposals on a carefully integrated set of strategic proposals, as follows:

- A series of tax breaks — worth Y100,000 million a year — is being planned to stimulate research and development in high technology. And MITI hopes to see in the next fiscal year tax credits, for as much as 50 per cent of increases in basic research expenditures, together with 50 per cent depreciation allowances for new research equipment and more favourable accounting procedures.

- The problem of the lack of venture capital for small companies is to be tackled. Tax-free loans for small companies in high technology, including software development companies, are proposed. MITI estimates there are some 3,000-5,000 small companies which could benefit from government support.

- MITI is planning to reorganize its own research efforts and to change its own rules

so that joint government-industry research, rather than only contract work, can take place inside government laboratories. MITI's numerous laboratories in Tsukuba Science City have superb modern facilities: access to MITI and its trained personnel will make joint research an attractive prospect for industry. And the resulting patents and their use by third parties will not be controlled by the government as at present but jointly by industry and government.

- MITI has an eye out for possible trade problems with other countries too. Next year it plans to set up a new institute, provisionally named the International Research Cooperation Japan Trust, modelled on the Humboldt Foundation of West Germany. Its aim is to avoid trade friction by promoting joint research projects. Postdoctoral research workers will be invited to Japan for one or two years at a time with the aid of fund which with luck will be provided in next year's budget and by donations from private industry.

Alun Anderson

Indian anaemia

Salt as iron vehicle

New Delhi

AFTER nearly eight years of research, the National Institute of Nutrition (NIN) in Hyderabad has found a practical measure to control iron deficiency anaemia (IDA), which afflicts more than half of India's rural population, mostly women. At NIN's recommendation, the Indian Government is taking steps to launch a national anaemia control programme based on iron fortified common salt (IFS) for the entire population.

The widespread prevalence of this anaemia is believed to be due to poor absorption of iron from the millet-based diet consumed in India, aggravated by the inadequate diet of the poor. Under Indian conditions, salt is an ideal vehicle for supplementing dietary iron. It is consumed by all and manufactured in only a very few places, where iron fortification facilities can easily be added.

NIN's choice of iron compound was ferrous sulphate, which is readily utilized by the body. The snag is that ferrous sulphate alters the colour of the salt. After a series of laboratory studies, NIN has devised a preparation that remains stable for eight months without change of colour or deterioration in quality. The formula: 3.5 mg of ferrous sulphate, 3.5 mg of orthophosphoric acid (which prevents colour change) and 5 mg of sodium acid sulphate (which improves the absorption of iron) in each gram of powdered salt.

Two long-term community studies have been carried out to demonstrate the effectiveness of IFS in preventing and controlling anaemia. In one study, IFS was given for 18 months to children of a residential school between 5 and 15 years of age. In another study, sponsored by UNICEF, IFS was supplied to rural households near Hyderabad and Calcutta and to urban dwellers in Madras city. In both studies, according to NIN, IFS was effective and free from side effects.

Even so, experts have pointed out that use of IFS for the entire population may not be without risk. It is feared that intake of extra iron over and above that derived from diet might increase the risk of infection. NIN, however, says that IFS consumption by normal people is safe and that, on the contrary, the available evidence shows that extra iron may actually improve 'immune status and other functional capacities of the population'.

The Government of India is at present also supplying iodized salt for the control of goitre in endemic areas in the sub-Himalayan belt. Since anaemia is also prevalent in that region, NIN is hoping to tackle both problems with salt fortified with both iron and iodide using a new formula now being evaluated. □

Industrial research

US tax credits a failure

Washington

WHAT seemed to be a straightforward scheme to encourage companies to put more money into research by offering special tax incentives has turned out to be a major disappointment, and it now seems likely that the four-year test of the programme will not be extended.

The United States thus appears ready to follow in the footsteps of Sweden, which recently abandoned its programme of tax credits for spending on research and development after a ten-year trial produced little evidence of any benefit to research activities in the private sector.

The US programme, which is costing the Treasury \$1,000-£1,500 million annually, allows companies to claim a 25 per cent credit for any expenditures on research and development that exceed the average of such expenditures for the preceding three years. The credit, which is applied directly against the company's tax liability, is in addition to the usual deduction against revenues allowed for all business expenses. In-house research costs plus 65 per cent of contracts (including contracts to universities) count towards the total.

The 1981 legislation that established the programme also allowed 'enhanced deductions' for companies donating state-of-the-art scientific equipment to educational institutions. Normally, companies can claim only their manufacturing costs when they make charitable donations of their products; the 1981 rule allows them to claim the full wholesale price so long as the equipment is new and working and the recipient is a college or university.

In practice, the results have fallen far short of the promise. A survey recently conducted by the General Accounting Office (GAO) found a significant number of service companies claiming the tax credit, generally for their 'research' efforts in developing software. An earlier survey by the Treasury Department found fast-food restaurants, bakeries, home builders, publishing houses and movie studios among them. These non-manufacturing and non-utility firms reported a suspiciously large growth in their research and development activities of 91 per cent between 1980, before the credit was introduced, and 1981.

Even among the manufacturing companies that claimed the bulk of the credits, there is little evidence that they were heavily influenced by it. In a non-scientific sampling of 86 companies, GAO found that slightly over one-third actually acknowledged that the credit had had no effect whatsoever on their research programme. Other statistics tend to confirm this conclusion more generally. According to Robert Eisner, an economist at Northwestern University, much of the

increase reported from 1980 to 1981 and subsequently was apparently the result of 'creative accounting' — redefining as research certain types of expenditures that were not considered as such in calculating the base rate of the previous three years.

Another factor behind the large amounts of credits being claimed was inflation. Companies were allowed to claim a credit for a nominal increase in research expenditures even when the increase only kept pace with inflation.

A survey conducted by Edwin Mansfield, an economist at the University of Pennsylvania, of almost 800 companies with research budgets greater than \$1 million showed that by their own estimates they would have spent only 1.2 per cent less in 1982 in the absence of the credits. Mansfield concluded that it was very unlikely that the tax-credits-induced research and development in 1983 exceeded \$683 million, while costing the Treasury \$1,000 million in lost revenue.

It comes as no surprise that the major industrial trade associations are staunchly behind a continuation of the programme. (A representative of the National Food Processors Association appeared before the House of Representatives Ways and Means Committee to spread the word that 'some of our innovative packaging and processing technologies might still be something in our future' were it not for the tax credits.)

A group called the National Coalition for Science and Technology (NCST), a self-styled 'science lobby' representing 1,000 scientists and engineers and about two dozen high-technology companies, also supports the programme. NCST's executive director, Marc Rosenberg, argues that four years is simply not enough time to expect a measurable effect from the programme, given the lengthy advance planning that goes into research decisions, and that all that has been done so far is to reward companies *ex post facto* for decisions taken before the credit was enacted. Uncertainty over the extension of the credit has also limited its effectiveness, he said. The credit is set to expire at the end of next year. The Senate has passed an extension; the House has balked, and is considered unlikely to act this year.

If a bill is passed, it is likely to include language that would tighten the definition of qualifying research to eliminate marketing research (which appears to have slipped in under the present rules) and emphasize technologically-based products and processes. But faced with the prospect of having to find additional revenue to reduce the \$200,000 million annual deficit, the Ways and Means Committee is not going to go easy next year on a bill that grants a \$1,000 million tax break to business.

Stephen Budiansky

US environmental protection

Agency pleads for less support

Washington

US FEDERAL agencies are not well known for complaining that they have been voted too much money. That, however, seems to be the attitude of the Environmental Protection Agency (EPA) to a recent vote in the House of Representatives to accelerate and expand the agency's "superfund" programme for clearing up toxic waste dumps. EPA says it would not be able to spend the House's proposed budget increase fast enough.

Superfund came into existence in 1980 to deal with chemical emergencies and, more controversially, the chronic problem of deteriorating waste dumps. There are at least 18,000 potentially hazardous sites known to EPA, and more than 500 have already been placed on a "national priority" list of sites posing a long-term threat to health or environment. The priority list could soon increase to more than 2,000 sites.

EPA's budget for the first five years of the superfund programme was \$1,600 million, but the House of Representatives voted overwhelmingly earlier this month to increase this more than sixfold, to a total of \$10,200 million for the next five years. But there were conditions. This apparent generosity was tied to a strict timetable for action: congressional critics of EPA maintain that the agency got off to a poor start on superfund and are fond of pointing out that only six clean-ups have been completed in the 4 years since the programme began. General support for EPA has decreased markedly, while many of its scientists, demoralized, have moved to jobs in academic institutions or industry.

The House of Representatives has now said it wants EPA to start on-site remedial action at 150 locations each year, as well as speeding up its study effort. Mr Lee M. Thomas, EPA's deputy administrator for solid waste and emergency response, has described these requirements as "unrealistic" in congressional testimony. Thomas estimated that the mandatory schedules alone would cost EPA \$7,400 million over the 5 years, with other provisions in the House bill increasing the cost still further. The result, according to Thomas, could be "the paradox of actually slowing down clean-up operations".

This stark prediction, perhaps unsurprisingly, cuts no ice with Representative James Florio (Democrat, New Jersey), who sponsored the House superfund reauthorization bill. Congressional analysts who support the House initiative say that EPA has deliberately exaggerated the cost of clean-ups to make a political point.

One controversial clause in the House bill that would have given the automatic right to sue in federal courts to those claiming to have been injured by toxic

chemicals was defeated by a narrow majority, much to the relief of the chemical industry. At present, claims against polluters have to go through the state courts first, where the standards of proof required may be more difficult for claimants to meet. Many state laws also impose time limits on claims which may make court action difficult if the alleged damage occurred many years before.

One revision to superfund that the house failed to take on board was a proposal, favoured by the chemical industry, to broaden the scope of its financing. Under the existing legislation, superfund is paid for largely by manufacturers and importers of 42 critical "feedstock" chemicals. This formula means that a high proportion of the cost falls to relatively few companies, while other historical polluters — such as

mining companies — get away scot-free. (Superfund is used only in an emergency and when a culprit cannot be found; the cost of clean-up operations rests with the polluter otherwise). The house did, however, retain the option of introducing a "waste end" tax — to be paid on waste produced. This form of taxation would, it is hoped, encourage more recycling and spread the cost of the programme in the way the chemical industry wants.

The Senate has not yet voted on its superfund reauthorization. Although some say it will follow the House lead during September, there is strong pressure on senators to wait until after the presidential election in November. The job is certainly not urgent, because the existing authorization will carry the programme through until the end of 1985. The administration fears that an overgenerous authorization will cause political damage, and is anxious to keep the issue out of sight for as long as possible.

Tim Beardsley

Albanian universities

Old problems found anew

POSTGRADUATE training at the Albanian university of Tirana is to be considerably expanded next term, with 50 courses as against 27 now, according to the Albanian press agency ATA. This appears to be a direct response to swingeing criticisms of Albania's higher education system by Ramiz Alia, chairman of the presidium of the People's Assembly, when he visited the university on 31 May. Since the purges of 1982, Alia has emerged as front runner in the succession to the ageing party leader Enver Hoxha, for whom he now increasingly deputizes. His speech at Tirana, the full text of which reached the West only at the beginning of August, took the form of a commentary on Hoxha's message to the university for its silver jubilee in 1982.

Higher education in Albania (Tirana University and a handful of agricultural, veterinary, pedagogic, physical culture and fine arts colleges) is, according to party doctrine, strongly geared to the needs of the economy; young people completing their studies are drafted to "serve the cause of the country and socialism wherever needed". Yet, according to Alia, higher education is frequently little more than a mechanical repetition by lecturers of textbook facts, while practical work and the training of students in technical dexterity is frequently neglected. Job-related training of undergraduates is "acute", and too little is being done to "discover and encourage new talent" and to develop students' "practical and creative spirit" and "technical intuition".

According to Alia, one of the main causes of this shortfall is the "mediocrity" and even "incompetence" of lecturers. During the past few years, he said, the party has been pressing for the replacement

of these unsatisfactory persons but found it difficult to identify them due to a "spirit of concession and compromise" still widespread in academic circles. Just how this cover-up has operated is not entirely clear since, according to Alia, the incompetents can be recognized by their poor standard of training, the long delays in their own progress up the academic ladder and the paucity of their publications. A simple check of *curricula vitae* and a scan of Albania's 30 or so learned journals would surely serve to locate the most flagrant offenders.

Tardiness in obtaining higher qualifications is, in Alia's opinion, the distinguishing mark of slackers who do not recognize that research is for higher education the "healthy and essential breath needed for progress". Some slackers, he complained, only defend one or two postgraduate theses, and then remain at that level. (Albania has a five-stage hierarchy of academic degrees and titles.)

At the same time, some of the ablest young graduates are immediately overburdened with teaching duties instead of being able to concentrate on their own postgraduate studies. The university and colleges, Alia urged, should make more "rational" use of such "*cadres*", since what the country most requires of them is "a high rapid and intensive qualification to the standard of contemporary science, . . . an integrated, scientific and professional training, . . . a solid mastering of basic foreign languages, a gradual but regular involvement in research activity both within and outside the higher schools and publication . . . from textbooks to genuine scientific monographs".

Vera Rich

Visa protests

Who will go to Prague?

JOHN Ziman, the British physicist, announces on p.619 his withdrawal from the Prague annual general meeting later this month of the European Physical Society (EPS). Ziman is protesting at the refusal of the Czechoslovak authorities to grant an entry visa to Czech emigré Frantisek Janouch (see *Nature*, 26 July, p.268), theorist and particle physicist, but it seems that Ziman will be protesting alone. He has circulated his letter to all other speakers at the Prague meeting, where he was to have given the Cecil Powell memorial lecture, but by last week neither he nor EPS had heard of any other withdrawals. The protest over the Janouch refusal seems destined to fizzle out, and EPS — says its president Godfrey Stafford — will not cancel the general council meeting on which Janouch was to sit. This says Ziman, is “scandalous and timorous”.

EPS, however, says that it would be impossible to reach agreement among the 90-member council, many of whom are East Europeans, to cancel the meeting, and that an attempt to do so might split the delicately balanced society in two. Anyway, says Stafford, EPS has been “no less successful” this time in arranging visas than at any other meeting. Only Janouch has been refused. Visas have been arranged, for example, for the whole Israeli delegation, although Czechoslovakia has no diplomatic relations with Israel, and even — after some difficulty — for a French supporter of the Charter 77 movement, outlawed in Czechoslovakia.

The main issue behind the dispute seems to be how far a “trans-bloc” scientific body such as EPS can go in pressing a case when the individual has both a scientific and a natural political interest in attending a meeting.

EPS also points out that visa refusals are not unknown in the West. At the last EPS annual general meeting at York, in Britain, an East German (Professor Bethge, of Halle) was refused a visa until the last moment, when it was too late for him to attend. At another EPS meeting at Grenoble, France refused a second East German (Professor Auth) a visa. EPS protested strongly on these occasions, but no explanations were given.

Ziman, however, says he would certainly protest at any visa refusal in any country. Moreover, he believes that Janouch is no more than outspoken about his mother country, and is not an activist. Janouch is, in fact, active in a Swedish committee supporting the Charter 77 movement, but believes such activity is “humanitarian and cultural” only, and not political.

The president of the Czechoslovak Academy of Sciences thinks otherwise. In a reply to Ziman's circulated letter, he expresses surprise that the British physicist

based his protest on “one-sided information” and that he had not tried to get information from the academy about Janouch. Explaining the successes with the Israelis and others, he goes on to say “the case of Janouch is rather different. He is a Czech emigré who has been, and is, working actively against the interests of our country, and you should not be surprised at the action of our authorities.”

Moreover, the letter goes on, “it would not be the first time that the activities of Janouch have threatened to split EPS”. This, presumably, refers to events of 1973, when the Czechoslovak Academy of Sciences was angling to host the 1975 EPS Congress, but was told by the then president of EPS, Professor H.B.G. Casimir of the Netherlands, that the society would be happy to accept the invitation, but that the situation of Professor Janouch could be a difficulty.

Dr Janouch had been dismissed from his job with the academy in 1970 on the grounds of “violating socialist order”, thus becoming the first victim in the academy of the purge “normalizing” the situation after the Prague spring. Although offered jobs in both Denmark and Sweden, Dr Janouch could not obtain an exit visa.

As a result of Casimir's intervention, the Czechoslovak Academy of Sciences cancelled its invitation for the conference. Shortly afterwards, Dr Janouch was informed, during an interrogation by the secret police, that it would be “more reasonable of him if he applied for a passport”, a suggestion reiterated by the academy authorities.

At the end of 1973, Dr Janouch was allowed to leave for Sweden, with a passport permitting him to return to Czechoslovakia within five years. In February 1975, however, he was formally deprived of his Czechoslovak citizenship, and when his appeal for reinstatement was rejected, had to apply for political asylum in Sweden. In 1979, he became a naturalized Swede.

Eighteen months ago, Janouch was elected to the general council of the EPS as one of the nine representatives of the “individual ordinary members”. Ironically, had it not been for Dr Janouch, there might never have been an East European presence in EPS, and the problem of “trans-bloc” conferences might never have arisen. During the negotiations in 1967-68 which led to the foundation of the society, it was Janouch, still at that time in good political standing, who persuaded the Soviet and East European physical societies to join.

Robert Walgate and Vera Rich

Particle physics

Der Spiegel makes trouble

DER SPIEGEL, West Germany's equivalent of *Time* magazine, threw the West German particle physics lobby into high gear last week with a scathing article by a British “professor of atomic physics” suggesting that the latest German particle accelerator,

raised by what physicists describe as the pure abuse of ParLOUR's article, its apparent authorship by an insider and the attention given it by publication in *Der Spiegel*. “It's crazy”, said a spokesman of the Deutsches Elektronen Synchrotron at Hamburg,

Sterbehilfe für eine Königin

Die teure und sinnlose Jagd nach den kleinsten Teilchen / Von Richard ParLOUR

Ist die moderne Teilchenphysik noch bezahlbar? Die Jagd nach den subatomaren Teilchen ist einer der höchstsubventionierten Wissenschaftszweige in den westlichen Industriestaaten. Allein der Speicherring, der gegenwärtig bei DESY in Hamburg im Bau ist,

kostet 250 Millionen Mark. Dieser Aufwand sei unsinnig, die Teilchenphysik eine sterbende Wissenschaft — so lautet die These von Richard ParLOUR, Professor für Atomphysik am Londoner University College, der den nachfolgenden Beitrag für den SPIEGEL verfaßt hat.

the DM6,000 million (£1,600 million) HERA, may be about as useful as the Egyptian pyramids.

The author, Richard ParLOUR, until recently a PhD student at University College, London, has touched a raw nerve. His view exactly matches that of West German environmentalists, the Greens, who now have political power through parliamentary representation won in the last general election.

To particle physicists, HERA will be a valuable complement to LEP, the electron-positron collider being built at the European centre for nuclear physics, CERN. HERA will ultimately be colliding leptons with quarks, providing results that should dovetail with one another.

West German tempers were particularly

where HERA is being built. “There's not a single argument in it, it's full of errors, and it ignores the past 20 years of great advances in particle physics.” Nevertheless, the article could do “enormous political damage” because of the resonance with the Greens' position and because “that's the kind of article politicians read”.

If the article had been published a few months ago, when the political decision to go ahead with HERA was being made, its effect could have been more damaging. Now, however, the sensitive and active West German high-energy physics community will be taking care not to over-react. There will be some discreet damage-limitation and some letters of protest to *Der Spiegel*.

Robert Walgate

European fast reactors

UK committee critical of pact

DESPITE reservations by a House of Commons Select Committee last month, Britain is to sign a major agreement with five other European nations jointly to develop a commercial fast-breeder nuclear reactor. That is the good news for the UK Atomic Energy Authority (UKAEA). The bad news comes in the government's review of the authority's future, commissioned in 1982 but not published.

The Select Committee on Energy, in its examination of Energy Research, Development and Demonstration in the United Kingdom (HMSO), singled out the fast-breeder programme for particular criticism. It has already cost £2,400 million and could cost £3,300 million more before a commercial reactor is in trouble-free operation.

The European collaborative effort entails the construction of three commercial reactors — in Britain, France and West Germany. (The other participants are Italy, the Netherlands and Belgium.) Sir Peter Hirsch, chairman of UKAEA, told the committee that it was necessary to build three reactors in order to satisfy divergent national licensing and operating procedures. The committee says that this seems to preclude common safety and design standards — a prime reason for such a cooperative venture in the first place.

According to the committee, cost savings for Britain would be negligible. Furthermore, it says, there is no evidence that Britain would face insurmountable technical problems if it opted to proceed with the fast-breeder programme alone. In the proposed agreement, Britain's reactor would be the last of the three to be built.

The Select Committee was also critical of the British thermonuclear fusion programme. Although scientific feasibility has always been in question, a major long-term commitment, denied to many less glamorous energy projects, has been made. Sir Hermann Bondi, a former chief scientist at the Department of Energy, is sceptical of the project, claiming that if a commercial fast-breeder reactor were developed, British long-term energy needs would be satiated and the fusion effort would no longer be necessary. On the other hand, if the breeder were to go, then the fusion programme should be scrapped also, as there is no evidence that commercial fusion will pose any less of a waste disposal problem.

These points tend to reinforce the Select Committee's view that nuclear projects are not subject to the stringent reviewing procedures of other energy projects. There is the feeling, fair or not, that, because of the enormous investment made in them, it is impossible to refuse demands for continued funds.

Marcus Chown

Satellite insurance

Shuttle to make good failure?

Washington

THE National Aeronautics and Space Administration (NASA) has formally announced a deal concluded with two insurance syndicates to retrieve and return to Earth the ill-fated Palapa B-2 communications satellite. The satellite belongs to the government of Indonesia and was launched into the wrong orbit last February. Negotiations still in progress could lead to a similar agreement to rescue Western Union Communications' Westar 6 satellite, launched at the same time. If successful, the mission would be a major publicity coup for NASA, which has taken a drubbing over the failure of the shuttle to live up to expectations.

The two satellites, although thought to be undamaged, have been useless since their launch from the shuttle. The failure of the rocket motors on the payload assist modules prevented the satellites from reaching geosynchronous orbit, and both are now in a parking orbit about 560 miles high. The two insurance syndicates involved, International Technology Underwriters of Washington, DC, and Merritt Syndicates Ltd, of London, have agreed to pay NASA \$4.8 million for the return of Palapa B-2. If (as expected) Westar 6 is eventually included in the package, the bill would be \$5.5 million — surprisingly little compared with the cost of a satellite launch, which can be up to \$70 million. The advantage to NASA will lie in the public demonstration that the shuttle can be used to return to Earth satellites not designed for retrieval. Hughes Communications will be paid \$2.5 million for having

involving the shuttle's remote manipulator arm, with the initial contact being made by a free-flying astronaut powered by backpack jets. The loading operation is planned to take about 4 hours per satellite. There could be complications in trying to maintain a suitable temperature, however. The Westar satellite would have to be picked up in an attitude that would allow it to get very cold during loading, and there is concern that this could threaten the operation.

The owners of the two satellites both received substantial insurance payouts when the satellites failed to reach the proper orbit in February. Under the agreement now reached to retrieve Palapa, ownership of the satellite would pass to the insurers before the operation started. The final agreement may include a provision for NASA to recondition the satellite once back on Earth. The price of a new HS 376 communications satellite of the Palapa and Westar 6 type is around \$35 million. If the rescue is successful, the insurers will thus be the first vendors on the used satellite market. The ownership issue is still not resolved in the case of the Westar satellite, but the deal to retrieve Palapa looks to be settled. The insurers' total loss on the failure in February has been estimated at \$75 million.

Tim Beardsley

Open University help for women

AFTER three successful years, a British scheme for providing financial assistance for women taking technical courses at the Open University is to be expanded. The scheme, run by the Manpower Services Commission and previously open only to women with a technical background, will now be extended to include those with no technical experience. Seventy places a year will be provided on the £42,000 a year scheme, compared with the hundred places offered in the past three years.

The technical courses cited for this special treatment, twenty-six in number, include electronics, telecommunications, engineering and computing. Successful applicants for a bursary will be able to embark immediately on one of these courses or, if required, take a first year "technology foundation course". There will be a preliminary "confidence boosting" weekend at Loughborough University, where students will be helped to identify a direction of study, their own weaknesses and the relevance of any previous job experience. Travel expenses to Loughborough and the annual week-long summer courses will be paid, leaving the bursary holders to contribute just £40 towards their yearly fees. **Marcus Chown**



"looked after" Palapa in orbit.

There are still some technical questions to be answered, however. The satellites will first be brought down under their own power to the shuttle's orbit at an altitude of about 200 miles. Both are at present rotating at 55 revolutions per minute and would have to be slowed down enough for an astronaut to approach, while still maintaining gyroscopic stability. The approach by the shuttle orbiter will be carefully contrived so as to avoid its thrusters being fired at the satellites.

If the approach is made successfully, the satellites will be transferred into the payload bay by a complex series of operations

Absentees from Prague meeting

SIR — I was shown an advanced copy of a letter from Professor Janouch shortly before its publication in *Nature* on 26 July (p.268). On 24 July I wrote to Dr Godfrey Stafford, President of the European Physical Society, as follows:

"The refusal of a visa to Professor Frantisek Janouch to participate in the 6th General Conference of the European Physical Society in Prague must be causing you grave concern. Janouch is not only a distinguished theoretical physicist, but also an elected member of the EPS Council, so that his enforced absence from this meeting would leave an intolerable gap in its scientific and professional proceedings.

I understand, from our telephone conversation the other day, that the EPS is taking this up with the Czechoslovak authorities, pointing out, no doubt, that their decision conflicts with the general principles on the organisation of international scientific meetings established by the International Council of Scientific Unions, and thus puts the whole meeting at risk. In support of any such representations, I should like to confirm to you in writing what I told you on the telephone — that I would not attend the Conference and give the Cecil Powell Memorial Lecture unless entry visas to Czechoslovakia are granted to all bona fide participants, including Frantisek Janouch.

I need not say that it would be a matter of deep regret for me personally not to take part in this meeting. It was a particular honour to be invited to give this distinguished lecture, and I was looking forward with great pleasure to revisiting the beautiful city of Prague, where I have many old friends. But such personal considerations are of little weight against the need to reaffirm the traditional principles of the international scientific community. I am sending copies of this letter to all the other invited speakers to the Conference, who must also have the same concern and may well be of the same opinion."

I did not make this letter public at the time, in order that further representations on this subject could be made to the Czechoslovak Government by both the EPS and the CS Academy of Sciences. I understand that such representations were indeed made, most forcibly, but it is now clear that they were not successful, and that the General Conference, together with the associated business meetings of the Council, of this distinguished international scientific society will be taking place in the enforced absence of one of its leading members. I trust that all participants in these

proceedings will understand and appreciate the grounds for my withdrawal, which are precisely as stated above.

J.M. ZIMAN

*Department of Social
and Economic Studies,
Imperial College,
London SW7 2PG, UK*

PhD glut in Japan

SIR — The special issue of *Nature* on science in India (12 April) attracted a flood of letters, whereas that on science in Japan (29 September 1983) apparently did not (see *Nature* 306, 220; 1983). The difference suggests that democracy, based on dialectics, is far more deeply entrenched among Indian than Japanese intellectuals. A test of this view of mine will be to see whether the response from Japan to the article by Alun Anderson (21 June) dealing with the burning issue of PhD glut, or "overdoctors" as they are called in Japan, will turn out to be a trickle or a deluge.

As Anderson rightly pointed out, the overdoctor (OD) problem is characterized by the indifference of the parties involved except the victims themselves. Japan's *laissez-faire* capitalism, the vicious competition which it engenders, in association with the hierarchical personal network which is almost the sole determinant of academic appointment and grant awards, is obviously the culprit.

It is therefore hard for those involved to raise their voices in a personally conspicuous way. Indeed, most of the activities of the OD campaign are actually run by graduate students — potentially ODs themselves — who will usually cease to be active once they have left postgraduate school for fear of offending their academic bosses.

The age distribution among academic staff in Japan will surely lead to deficiencies in Japanese science in a decade or so, but the lack of postdoctoral fellowships in Japan seems to worry overseas observers more seriously than the Japanese. In fact, I know of no article on the Japanese OD issues which is as lucid and to the point as Anderson's.

In reality, the problem is intimately linked with the university structure in Japan¹. Some ambitious scientists in provincial national universities, for example, are currently struggling to open doctor courses in their universities at this very time, but are frustrated by the university and the divide-and-rule tactics of the government. The OD problem is generated by elitism of the major national universities.

I cannot help feeling that the whole situation is alarmingly similar to that which led to the student protests in the University of Tokyo in 1968². At the beginning, the students identified themselves as victims within that arch-elitist university. When the protest gathered momentum towards the end of 1968, however, the students

realized that, after all, their protest was a cosy way to join the elites of which they were complaining, hence the famous slogan of "self-denial" launched at that time. For a time, the defects of the university system in Japan were fiercely debated but the outcome was that reform was shelved with the "normalization" of the universities called for after the ugly and violent campus confrontations.

One of the features of that time was that the insurgent students did not seek solidarity with citizens in general, in contrast with what happened in Paris in May 1968. Unless the OD can surmount the barrier between town and gown, that campaign may also fail.

As for immediate remedies, government and industry should be urged to create a huge fund for postdoctoral fellowships (instead of EXPO etc.), which however should be open to overseas applicants as well, on a fair, competitive basis, with less or even no emphasis on proficiency in Japanese. University professors should be given points for their record of collaborating with ODs as co-authors of published papers. Also science shops, quite successful and widespread in the Netherlands (*Nature*, 7 June) to provide enquiring citizens with services in return for fees, should be considered by ODs as an alternative source of income.

ATUHIRO SIBATANI

*Department of Mathematics,
Kyoto University,
Kyoto 606, Japan
(on leave from CSIRO,
Sydney, Australia)*

1. Nihon Kagakusha Kaigi (ed.) *Overdoctor Mondai (PhD Glut in Japan)* (Shoten, Tokyo, 1983).
2. Sibatani, A. *Nature* 240, 191-193 (1972).

SIR — Your interesting news item (*Nature* 21 June, p.659) on "overdoctors" (ODs) in Japan misses an important point. PhD students are essential in a Japanese research group not only because of the heavy teaching load of the staff but because of the need for someone to supervise the large numbers (often 50 per cent or more) of final-year undergraduates who stay on to do the two-year MSc course.

Japanese industry wants people with a first degree, prefers those with an MSc, but regards PhDs with some suspicion. If industry, recognizing the supervisory role of a PhD student during his training as well as his enhanced technical skills, were to welcome PhDs, the OD problem would be much reduced. British industry has in recent years transformed its attitude to the PhD degree, and there now exists a variety of non-university options for a graduate student who cannot, or does not want to, pursue a research career.

A.J. MACKINTOSH

*Cavendish Laboratories,
Madingley Road,
Cambridge CB3 0HE, UK*

Two views on UK research

SIR — We were heartened to see your leading article "Dead-end for British research" (*Nature* 26 July, p.261), highlighting the fact that imaginative research is being killed off in this country. We support your plea to the government to act quickly, but not the contention that it alone is the culprit. We were astonished that no mention was made of the fact that the people actually doing the research, respected senior and junior scientists throughout the country, saw the warning signs as long as ten years ago, which resulted in the formation of the Association of Researchers in Medicine and Science (ARMS) in 1978. The publicity in the scientific and lay press at the time of our lobby of Parliament in 1979, highlighted our concern in urging a major review of the whole research and development sector, and pointed out the irretrievable losses and delays to scientific and technological progress that would result if nothing was done (see *Nature* 282, 554; 1979).

The basic problem lies not with the government but in the fixed attitudes in the research councils and universities as to how scientific research should be done. We believe that for research to prosper, at least some of it must become more professional, and not be exclusively the amateur part-time adjunct to teaching that it has historically been. This will necessitate establishment of research careers and recognition that the new graduate entering any research sphere needs training followed by probation. Recognized talent should then be rewarded with a career supervised by a body of professional researchers — as in any profession.

Despite our protests that today's young were more realistic and would turn away from research, to the detriment of research and the universities, nothing has been done. In fact, quite the reverse. The number of staff on short-term age-restricted grants has risen to 23 per cent of the university academic work-force, and the advice of the Advisory board for the Research Councils (ABRC) in July 1983 was to extend these contracts not curtail them. ABRC proposes to increase the number of short-term posts in units where a career structure actually exists in order, among other things, to "employ experts in some special or new technique to train existing staff" (grants for not more than 3 years, and a maximum of 5 years). "Post-doctoral fellowships to attract scientists (20-40) at the height of their scientific prowess to interact with existing staff" were also recommended (but to terminate in 3 years — positively without extension or renewal). All the above would be subject to the waiving of redundancy/unfair dismissal rights.

Are these attractive proposals? Why should the highly experienced, or the young, be used to "maintain the vigour and stimulus" of an outdated system, happy to discard them in 3 years or less? It shows no consideration for the people who will do the research, short or long-term, let alone for the quality of research such a system should produce. This smacks more of the treatment meted out by bullies to the nineteenth century chimney-sweep — not the advanced thinking vital for the sophisticated technology of the twenty-first century one would expect from a board controlling the nation's research.

Small wonder that with the prospect that at thirty (now much older than their fellow graduates who opted for a job with some form of career advancement) they will be unemployable, the young are now scorning not only research but also science. Our prediction in 1979 that such continuing policies would confirm the United Kingdom's position as a third-rate economic and scientific power, with a complement of first-rate scientists and scientific opportunities going to waste, has really come to pass.

Our proposals would attract back the young graduate, but urgent action is also needed to prevent the irreparable loss of key senior personnel. Implementation of our proposals would involve a change in the dual support system. Since it has completely broken down as a result of government cuts, now is the time for action.

H. ANNE SIMMONDS
*Association of Researchers
in Medicine and Science Limited,
c/o Clinical Research Laboratories,
Guy's Hospital,
London SE1 9RT, UK*

SIR — Your leading article of 26 July is ill-judged, ill-informed and opinionated. It is one thing to moan about the financial problems of British universities and research councils, but it is quite another to belittle the efforts and achievements of British scientists in teaching and research. That you can find nothing in recent British research to lift your spirit or stretch your imagination reflects, sadly, the state of your spirit and your lack of imagination (notwithstanding your doubtful clairvoyance in the matter of Nobel prizes) rather than any shortcomings on our part. You say that we should take steps to defend ourselves against mediocrity, but surely you must appreciate that almost all scientific research is "mediocre" insofar as it is unworthy of banner headings in *Nature* and unlikely to be rewarded with a Nobel prize.

A greater threat to British science than any supposed mediocrity within the research community is the onslaught of irrationality from without; for instance, you parody the British PhD system as "consisting of giving a young person a research problem and telling him [not her?] to come back when it is solved". If by this

you mean that British research students are encouraged to think for themselves then it is hardly, as you say, "amateurish"; or do all those foreign professionals whose work you so admire get someone else to do their thinking? Have you ever heard of Brian Josephson, who did rather well as an "amateur" British research student by thinking for himself.

Perhaps you should try this thinking lark for yourself the next time you compose an editorial rather than just stringing together clichés. It takes some practice, but you might find it produces a more cogent argument, and one which focuses on the real problems affecting science.

NEIL THOMAS

*Department of Physics,
University of Birmingham,
PO Box 363,
Birmingham B15 2TT, UK*

Reform needed

SIR — Somewhere in the back issues of *Nature* is the observation that, while in the nineteenth century the creative scientist could possibly expect starvation, in this century he is more likely to be buried in a mass of trivia. Unfortunately, public support for science also has more serious consequences, some of which are already evident. On the one hand, relatively large funding has attracted power-hungry people, (operators, grantsmen) into research, and on the other administration of these funds invokes the complementary bureaucratic ethos of playing it safe and other games of expediency and evasion. Consequently peer review, in principle the most effective method of evaluation, is seriously distorted both by incompetence and by unidentified conflicts of interest, in particular undue influence of operator types. Decisions reached in this way are not subject to any form of appeal or accountability. Bureaucrats also seem remarkably insensitive to ethical issues, especially when they reveal embarrassingly poor judgement on their part.

A number of reforms are needed: (1) An effective appeals process; (2) accountability of review and award panels; (3) a precisely defined and strictly enforced code of ethics; (4) limitation of grant awards to those in which the grantee is performing directly on a day to day basis (that is, no absentee "principal investigators").

It should be recognized that administration of science is a different undertaking from that of other public enterprises, since the primary unit is the individual investigator who should not be forced into the position of becoming an entrepreneur in order to survive. On the contrary it is the job of the administrator to protect him from political currents which have little to do with the practice of science.

M.C. GOODALL

*University of Alabama,
Birmingham, Alabama 35486, USA*

Widespread after-effects of nuclear war

from Edward Teller

Radioactive fallout and depletion of the ozone layer, once believed catastrophic consequences of nuclear war, are now proved unimportant in comparison to immediate war damage. Today, "nuclear winter" is claimed to have apocalyptic effects. Uncertainties in massive smoke production and in meteorological phenomena give reason to doubt this conclusion.

CORRECT identification of danger is essential to intelligent action. Scientific knowledge of the after-effects of a nuclear war — fallout, ozone layer depletion and climatic effects — is of great importance in making political decisions. Early fears concerning two of these, fallout and ozone depletion, turned out to be exaggerated. A third, the fear of climatic changes due to smoke and dust, is under active discussion now. The study by Turco *et al.* is an early contribution to the question¹. The preliminary and uncertain nature of their findings deserves clarification.

Fallout

The consequences of war-engendered fallout on non-combatant nations were the first to receive widespread popular attention. During the 1956 Presidential campaign, Adlai Stevenson raised the problem of worldwide fallout in an arresting manner. In 1957, a book, *On the Beach*, and its subsequent movie version popularized and perpetuated the belief that the effects of fallout would endanger, and probably extinguish, all human life².

In 1975, the US National Academy of Sciences issued a report in which the estimates of war-produced worldwide fallout were so low as to have insignificant biological consequences³. However, estimates used in the study were scaled from larger nuclear weapons tests and, therefore, were not representative. Actual weapons tests avoided atmospheric conditions where rainfall was expected, and were not carried out in middle latitudes. Because of their size, a larger fraction of radioactive products from the tests was injected into the stratosphere than would be from current weapons and, because of the longer residence time in the stratosphere, the radioactivity of fallout from a nuclear war was underestimated.

The best current calculations, made at Lawrence Livermore National Laboratory (LLNL), suggest that global-scale fallout will expose people in the non-belligerent countries in the Northern 30 to 70 degree latitudes to an average external dose of 20 rem — a total dose delivered over 50 years⁴.

It is important to note that this estimate holds on the average. Vagaries of weather, particularly of rainfall, can concentrate fallout in large but limited areas called hot spots, where the total dose values in non-belligerent countries could be as high as 250 rems⁴. (The assignment of these doses does not take into account the shielding afforded by houses, which would reduce the radiation received by up to one-half.)

A given radiation dose received over a long period is generally less harmful than the same radiation dose received in a short time, because of continual biological repair^{5,6}. The formula for slow dose rates in current use suggests that 90 percent of the damage is repaired at the rate of 2.5 percent per day⁵. An acute dose of 100 rem or less does not cause appreciable radiation sickness. When the biological recovery formula is applied, a dose as high as 250 rem from worldwide fallout is built up at a rate so slow that it would not produce significant casualties.

Inclusion of radioactive elements in the food chains might, on the average, increase these estimates by a factor of two⁴. Simple precautions based on generally held knowledge about fallout, such as that offered in public schools in the Soviet Union, could sharply limit this danger. The contrast between the long-term danger of fallout to the worldwide population and the great immediate dangers in combatant countries is extremely marked.

The worst case of fallout would occur if all nuclear reactors and all used reactor fuel elements were subjected to explosions violent enough to disperse their radioactive contents. Under these conditions, the 20 rem average dose and the 250 rem "hot spot" dose may need to be multiplied by a factor of about eight⁴. Fallout in the non-belligerent northern hemisphere countries could then be a considerable danger. Vaporization and dispersion of reactor radioactivity is not easily accomplished nor does it appear to serve any important military purpose. It is therefore unlikely to occur. Early underground disposal of wastes would greatly diminish this danger.

Thus, there is almost no likelihood that fallout outside the belligerent nations (with the exception of regions close to the belligerents) would have a major damaging effect. Even in the belligerent countries,

fallout would have less important effects than the immediate effects of fire and blast if the most elementary civil defence measures were exercised. Fallout is an example of a consequence that further study has shown to be far less important than originally feared.

Ozone layer damage

In the 1970s, there was concern about the possible effects of a nuclear conflict on the ozone layer^{3,7}. The small amounts of ozone present in the stratosphere play an important role in shielding the Earth from ultraviolet radiation. Large nuclear explosions propel the contents of their fireballs, including the molecules NO and NO₂ (collectively referred to as NO_x), into the high atmosphere where they react with ozone, eventually converting ozone to oxygen molecules (O₂).

If injected in sufficient quantity at high altitude, the NO_x could deplete the ozone layer significantly. Ozone molecules are steadily regenerated by solar radiation, but the NO_x residue from large nuclear explosions could reduce the effective thickness of the ozone layer by an average of 30 to 40 percent for a period of a year or two^{8,9}. (For purposes of comparison, the effective thickness of the ozone layer during summer is about 20 percent less over Miami than Seattle.) Complete recovery would take several more years⁹.

Assuming the explosion of thousands of megaton weapons, the calculated diminution of the ozone layer would give rise to increased serious sunburn immediately and to an increased incidence of skin cancer over an extended time. These primary effects, like the possibility of eye damage, could be countered with simple precautions. A secondary concern is that ozone depletion and the consequent increase in ultraviolet radiation would destroy ultraviolet-sensitive plants or plankton species, and would harm early springtime plant growth, with important effects on ecosystem viability¹⁰.

Since, in middle latitudes, observed natural variations of the ozone column are comparable with the predicted effects of a nuclear war with megaton weapons, the changes are not likely to be highly significant¹¹. The most pertinent consideration, however, is that the present US

Edward Teller is at Lawrence Livermore National Laboratory, Box 808, Livermore, California 94550, USA.

arsenal has few nuclear explosives with the great yield that is a necessary precondition for depletion of the ozone layer. The same trend toward smaller, more accurately aimed and less costly weapons also has been adopted with some delay by the Soviet Union.

In conclusion, problems related to a weakening of the ozone layer seem manageable, and with the rapid elimination of weapons large enough to cause ozone layer damage, the probability of serious ozone layer depletion is dropping significantly.

Effects of dust

In the mid 1960s, the climatic effect of dust raised by multiple nuclear blasts became a concern at Lawrence Livermore National Laboratory. (Unfortunately, this early work by R.G. Gutmacher and G.H. Higgins is still classified.) Dust in the post-attack atmosphere would tend to reduce solar radiation received at the surface of the Earth. If this reduction were sufficiently large and prolonged, the hemispheric temperature would be lowered.

The effects produced by blast-lofted dust from nuclear explosions can be expected to be analogous to those produced by dust spewed into the atmosphere by volcanic eruptions. Annual growth rings in bristlecone pine trees show the effects of early frost, and these changes are correlated with large volcanic eruptions that distributed dust in the upper atmosphere¹². Comparing the historic volcanic events with annual growth rings offers a rough measure of the relationship between dust and worldwide temperature³.

The quantity of small particles in the dust raised by a large scale war is comparable with that produced in the largest volcanic eruptions. Recent estimates of the mass of small dust particles that would be lofted into the stratosphere in the event of a nuclear war are based on scanty information. Review and analysis of old air filter samples and recent tests involving large amounts of high explosives may provide a more reliable basis for calculations.

The most probable conclusion is that the atmospheric effects from war-generated dust would be noticeable but by no means severe on a hemispheric scale.

Effects of smoke

In 1982, Crutzen and Birks proposed that smoke particles from forest fires produced by a nuclear war would give rise to a much more serious effect than dust. Smoke absorbs sunlight as well as scattering it¹³. In 1983, Turco *et al.* argued that smoke generated by fires in bombed cities or oil refineries would produce even greater effects than those from the forest fires¹. In discussing these results, some scientists have created the impression, aided by a short movie depicting the postulated effects, that smoke from fires ignited by nuclear explosions and, to a lesser extent, the dust raised by a nuclear exchange would

bring about what is now known as "nuclear winter"¹⁴.

The theory of a severe nuclear winter depends on the assumption that a large quantity of smoke will be emitted by burning cities and forests and will be distributed in the troposphere. Such smoke-laden air, heated by sunlight, could raise the temperature at the tropopause from approximately -50° to -60° C to about 5° C. The surface temperatures of the continents in the Northern Hemisphere between the latitudes of 30° and 70° might drop to about -30° C because of the absorption and scattering effect of the smoke.

A large quantity of smoke at high altitudes would produce an extremely massive inversion, a state in which the temperature increases, rather than decreases, with altitude. In regions of inversion, rising air masses lose their buoyancy, so water condensation and rain is suppressed. Thus, smoke would persist in the upper atmosphere, shielding the Earth's surface from sunlight. This situation could persist for long periods of time, and the resulting temperature change would damage all life in the underlying region. The argument, as continued by Sagan, holds that smoke may well be distributed over the entire globe and that the survival of the human race could be endangered. Although the findings of Turco *et al.* are inconclusive, Sagan calls them scientifically robust¹⁴.

Examination of the possibility of nuclear winter differs from ascertaining fallout and ozone layer depletion because it depends on meteorological phenomena that involve much more detailed and complex considerations and calculations. Several important meteorological processes are inadequately understood, and the bases for smoke estimation are very uncertain. Computer modelling sufficiently detailed and refined to handle all the various pertinent factors is not yet available.

It would take about two weeks for the smoke to disperse over wide areas. Turco *et al.* do not consider¹ in sufficient detail the abnormal mechanisms that may be operating during this period which could sharply decrease the amount of smoke available. There is inadequate quantitative allowance for the effects of several major meteorological phenomena in the prediction of the nuclear winter, and the bases used for smoke estimation are most uncertain.

The average tropospheric residence time of fine dust particles, such as those containing radioactivity, varies between one and three weeks⁴. The average residence time of water vapour in the atmosphere is little longer than a week¹⁵. Studies of the continuing discharge rates of man-made and natural smoke, and observations of the average amount of smoke found in the atmosphere, suggest that smoke may have an atmospheric residence time of a week or less¹⁶.

Under all atmospheric conditions, water

is the agent largely responsible for cleansing the atmosphere of smoke. Turco *et al.* estimate that 225 million tons of smoke particles (2.25×10^{14} grams) would be emitted in their 5,000 megaton baseline war¹. A comparison between this weight of smoke and the weight of the water (droplets and vapour) in the northern latitudes between 30° and 70° shows that the water outweighs the smoke by a factor of ten thousand. The average lifetime of a water molecule in the atmosphere is little more than one week. So, in ten days, before the widespread smoke-associated cooling effect is established, a mass of water at least a few thousand times greater than the total mass of war-engendered smoke will rain out of the atmosphere. Turco *et al.* estimate that a factor of three or four of the smoke (depending on detailed conditions) will be removed by rainout in the fourteen day period.

The highly relevant question of how much smoke will remain aloft after two weeks of rainout is completely unanswered because it depends on the relative altitudes of water and smoke, on alterations in scavenging processes, on latitudinal variations and on various other factors. Two studies suggest specific cleansing factors.

The capping clouds which regularly appear in the smoke plumes above big fires are produced by condensed moisture that forms droplets too small to develop into rain. In one observation made for a US Environmental Protection Agency 1983 report (preprint, University of Washington, Seattle), L.F. Radke *et al.* measured the particle size distribution for smoke following two different paths, one through the cloud and one below it. A comparison of these two size distributions can be interpreted to say that the cloud reduced by a factor 10 the number of small smoke particles between one-tenth and one-half micron in size. The effect would be large because most of the particles would pass through the cloud. Unfortunately, this experiment has not been repeated for lack of funds. More experiments are now being planned.

Changes produced by the capping cloud may result from each of the water droplets trapping many small smoke particles. Upon subsequent evaporation of the water droplets, these particles are agglomerated into bigger units. In addition, Hanel observed that aerosols increase in size if they are exposed to air in which the relative humidity exceeds 90 per cent¹⁷. Increasing the sizes of the particles reduces their number in the range where they are most effective in absorbing and scattering the sunlight. Larger smoke particles also rain out and fall more rapidly. Thus, at the very origin of smoke generation, the particles already may be prepared for more rapid deposition.

The nuclear winter mechanism for the abrupt cooling of the Earth's surface and the lower troposphere is effective only over

land. Over the oceans, the surface water is continually supplied with heat from below. For periods of many months, the temperature at the ocean surface would remain essentially unchanged.

The Turco *et al.* model for predicting nuclear winter did not include the effects of oceans or winds in a quantitative manner¹. The National Center for Atmospheric Research (NCAR) uses a somewhat more realistic model of the atmosphere in that it includes oceans and winds¹⁸. In the NCAR model, the extent of the temperature reduction is lower by a factor two to ten (depending on the season) than in the Turco *et al.* study.

However, the NCAR model still postulates that the injected smoke layer appears instantaneously at all temperate zone longitudes. Furthermore, the NCAR model postulates that solar radiation is absorbed by the smoke layer as though the layer were fixed in its longitudinal position and concentration. This postulate ignores the motion of the smoke produced by convection processes related to natural induced air flow. Finally, in the absence of a usable model, the NCAR study postponed a consideration of the most important factor of rainout. While the NCAR study is the most elaborate calculation yet produced, it does not provide estimates that can be considered final, even in an approximate sense.

If a nuclear winter were to begin, unusually great temperature differences would be established between continents and oceans. This results in increased storms along the eastern coasts of the continents, with consequent mixing of air from high and low levels of the troposphere. Air from lower altitudes has a higher moisture content which is conducive to rainout. The storms are mentioned by Turco *et al.* but not taken into account in the calculation of particle scavenging¹.

The warm wet air over the oceans might rise only to a fraction of the postulated stable altitude of the smoke layer. It is well-known, however, that storms and associated turbulent air quite often extend well above the region of the original source of the storm and may transport moisture up to the tropopause. If the rain removed smoke from one area while leaving it untouched in another, the smoke layer would be patchy. The obstruction of light is most effective when smoke is well spread. The hemispheric light reduction would decrease as patchiness increased.

The patchiness of smoke, in turn, could bring about further large temperature differences which will further drive the atmospheric processes that rain out more smoke. In this case, rainout of large smoke layers might proceed in an accelerating fashion.

Some smoke particles, particularly those from oil fires, absorb light strongly. At first these particles may be water-repellent, as Turco *et al.* suggest. This property changes after relatively brief exposure (24 hours) to



Capping cloud created by a forest fire at Reno, Nevada in July 1984. The photograph shows the smoke plume, the entrainment of the smoke into the central portion of the capping cloud, and the de-entrainment of the smoke aerosol from the lower portions of the capping cloud. (Photo by John Hallett of the University of Nevada Desert Research Institute.)

natural oxidizing agents in the atmosphere. Such agents are likely to persist, at the very least, in the relatively clear patches intermixed with the smoke. Pure graphite is water repellent, but smoke usually consists of graphitic particles covered by organic material. Alternatively, large particles would act as condensation nuclei even if hydrophobic. Radke's observation suggests that wood fires produce water-attracting particles even initially. Therefore, there is little reason to expect that rainout of all types of smoke particles would not occur.

In conclusion, there is no obvious reason why large quantities of smoke particles should have a long residence time in the atmosphere. The calculations associated with the nuclear winter do not include major cleansing effects of water vapour which are themselves smoke-induced or the influence of the oceans and winds during the time needed for spreading the smoke world-wide.

A recent detailed study by A.A. Broyles, based on observations of forest fires and considering the proper modifications pertinent to a discussion of burning cities, proposes that the upper limit of smoke that might be emitted in a nuclear war is 360 million tons while the lower limit is 15 million tons¹⁹. These limits are not rigorous; somewhat higher or lower figures are possible. However, the Turco *et al.* estimate for their baseline war producing 225 million tons is in the upper range.

The uncertainties associated with the amount of smoke that would be emitted are

extremely numerous. The ratio of wood burned to smoke emitted depends on many factors, including moisture content, the temperature of the air surrounding the fire, the manner in which the wood is packed and the accessibility of oxygen. The last two factors depend respectively on the size of the object burning and on the wind velocity.

Estimates of the quantity of smoke generated are based on assumption that the amount of smoke is proportional to the ratio of the weight of the smoke emitted to the weight of the fuel burned. The measured smoke emission factor presented in one forestry report ranges from 0.0025 to 0.063 for wood²⁰. Measurements made at Georgia Technical Institute give values ranging from less than 0.01 to 0.2 for plastics²¹. There is some possibility that large urban fires may provide sufficient heating reinforcement among burning components to consume even more of the smoke that is indicated by these numbers so that even lower values are applicable.

Estimates of the quantity of smoke also depend on the area burned, which in turn depends on many factors including the radiant energy flux required to produce ignition, atmospheric transmittance and topography. Estimates of the density of fuel in the areas burned also involve a large factor of uncertainty. The amount of wood in forests is usually considered to vary between 0.5 and 10 kilograms per square metre²²; the corresponding values are between 20 and 200 kg m⁻² in cities²³⁻²⁵.

In cities with the higher fuel density

values, fire storms are possible. Fire storms which loft smoke to very high altitudes are relatively rare events, apparently dependent on the presence of dense fuel and particular meteorological conditions under which practically all available oxygen is consumed. The possibility that firestorms may loft smoke into the stratosphere cannot be excluded, but has never been observed. Turco *et al.* assume that only 5 percent of the smoke produced is due to firestorms¹. A great deal of variability exists, and at present, there is no reliable basis for large scale estimates of smoke emission from burning forests and cities.

Smoke generated by burning petroleum products, rubber and plastics tends to absorb light more strongly than the smoke from burning wood. The predominant fuel in cities, at least 95 percent, seems to be wood and paper products^{23,24}. The strength of light absorption of city smoke is critically dependent on the proportion of petroleum products, plastics and rubber present. This uncertainty strongly influences possible temperature changes.

Petroleum refineries could provide a total of 150 million tons (1.5×10^{14} grams) of fuel in the United States²⁶. The weight of smoke would be about 15 million tons. Much of it would contain sulphur oxides, and less than 10 per cent is apt to consist of strongly absorbing carbonaceous particles. Oil found or stored underground is not apt to be ignited. Thus, the strongly absorbing carbon-containing smoke from this source is small in quantity.

Another major uncertainty connected with the properties of smoke is based on an inability to predict how a war would be conducted. If only purely military targets were attacked with relatively small bombs, the amount of smoke generated would be much lower than that required for a nuclear winter, and the argument for the unavailability of a nuclear winter would lose all force. Because of the trend toward more accurate, lower-yield missiles, indiscriminate damage — the source of great quantities of smoke — is far less likely. The deployment of anti-ballistic missile defenses would reduce further fire and smoke.

Conclusions

During the two-week period required to establish the full effects of a nuclear winter, the water in the atmosphere, ten thousand times greater in weight than the postulated emitted smoke, most probably will rain out. This alone could reduce the smoke content at low altitudes manyfold. Other weather phenomena, enhanced by the patchiness of the smoke and the temperature differences between land and ocean, will serve to bring smoke from the upper levels of the troposphere to the lower, whence it can be removed. The immediate

coagulation effect of capping clouds on the smoke is unknown, but it might reduce the small smoke particles emitted (which are the most dangerous) tenfold. The targets that would be attacked are uncertain. The proportions of highly light-absorbent smoke that would be emitted are also unknown. Complex factors such as fuel density, moisture content, air temperature and other properties related to smoke emission make current estimates unreliable. The associated effects of each of these uncertainties on temperature change tend to decrease the amount of smoke retained in the atmosphere. The effects of these uncertainties are large enough that revising a few of the smoke estimates and including a few meteorological effects invalidates the occurrence of a severe nuclear winter.

Given the uncertainties and omissions in the theory on which nuclear winter is based, the concept of a severe climatic change must be considered dubious rather than robust. Nonetheless, the possibility of nuclear winter has not been excluded.

The probability that war-engendered smoke and dust would reduce the temperature by at least one-tenth the number of degrees that Turco *et al.* predict, is obviously much greater than the probability of the occurrence of extreme temperature changes. A decrease of 5° or 6° C between northern latitudes 30° to 70° during summer, rather than the predicted 50° or 60° C, could still lead to widespread failure of harvests and famine.

Therefore, there is every reason to undertake the difficult task of arriving at more realistic predictions. More funds should be allocated for experiments to sample fire plumes and to determine how emissions scale with the size and type of fire. Greater amounts of money, carefully spent on atmospheric modelling and experiments, would accelerate resolution

of the basic questions regarding nuclear winter.

From Biblical times to the present written records testify to times of food shortage from causes ranging from volcanic activity to mankind's folly. The amount of suffering that would be produced today by the failure of a single year's harvest is horrifying to consider. Extensive food storage could decrease greatly the risk of famine and could be easily accomplished in the United States, and in many other nations. The usefulness of such storage is unquestionable.

Highly speculative theories of worldwide destruction — even the end of life on Earth — used as a call for a particular kind of political action serve neither the good reputation of science nor dispassionate political thought. Informing the general public about the reliability, unreliability or incomplete nature of such studies seems an important responsibility, for only then can decisions affecting the well-being of Western societies be made on an intelligent basis.

I thank the scientists in the LLNL Atmospheric Sciences Division — Joseph B. Knox, Michael C. MacCracken, Frederick M. Luther, and Joyce E. Penner — for the research on which my main conclusions are based and for their comments on this paper. Particular thanks are due Professor Arthur A. Broyles, University of Florida. Finally I thank the Hoover Institution, Stanford, California, for continuing editorial support. This work was performed under the auspices of the US Department of Energy under contract W-7405-Eng-48. The views and opinions are nevertheless those of the author alone, and the usual US government disclaimer of legal responsibility, explicit or implied, should be understood as accompanying this article. □

1. Turco, O.R., Toon, P.B., Ackerman, T.P., Pollack, T.P. & Sagan, C. *Science* **222**, 1283-1292 (1983).
2. Shute (Norway), N. *On the Beach* (Morrow, New York, 1957).
3. *Long-Term Worldwide Effects of Multiple Nuclear-Weapons Detonations* (National Academy of Science, Washington, D.C., 1975).
4. Knox, J.B. *Global Scale Deposition of Radioactivity from a Large Scale Exchange* (LLNL Report UCRL-89907, 1983).
5. National Committee on Radiological Protection & Measurement, Report 29 (Government Printing Office, Washington, D.C., 1962).
6. *The Medical Effects of Nuclear War* (British Medical Association, London and Wiley, New York, 1983).
7. MacCracken, M.C. & Chang, J.S. *A Preliminary Study of the Potential Chemical and Climatic Effects of Atmospheric Nuclear Explosions* (LLNL Report UCRL-51653, 1975).
8. Chang, J.S., Duewer, W.H. & Wuebbles, D.J. *J. geophys. Res.* **84**, 1755 (1979).
9. Luther, F.M. *Nuclear War: Short Term Chemical and Radiative Effects of Stratospheric Injections* (LLNL Report UCRL-89957 1983).
10. Caldwell, M.M. *BioScience* **29**, 520 (1979).
11. *The Effects of Nuclear War* (Congressional Office of Technical Assessment, Washington, D.C., 1979).
12. LaMarche, V.C. Jr & Hirschboeck, K.K. *Nature* **307**, 121 (1984).
13. Crutzen, P.J. & Birks, J.W. *Ambio* **11**, 115, (1982).
14. Sagan, C. *Foreign Affairs* **62**, 257-292, (Winter 1983).
15. Sellers, W.D. *Phys. Climat.* (University of Chicago, 1965).
16. Sciler, W. & Crutzen, P.J. *Clim. Change* **2**, 207-247 (1980).
17. Hanel, G. *Atmos. Envir.* **16**, 2703-2706 (1982).
18. Covey, C., Schneider, S.H. & Thompson, S.L. *Nature* **308**, 21-25 (1984).
19. Broyles, A.A. *Am. J. Phys.* (in the press).
20. Tangren, C.D., McMahon, C.K., Ryan, P.W. & Zinn, B.T. in *Southern Forestry Smoke Management Guidebook* USDA Forest Services Southeastern Forest Experiment Station, Asheville, North Carolina, 1976).
21. Bankston, C.P., Cassanova, R.A., Powell, E.A. & Zinn, B.T. *Review of Smoke Particulate Properties Data for Burning Natural and Synthetic Materials* (NTIS PB 292066, 1978).
22. Brown, A.A. & Davis, K.P. *Forest Fire Control and Use* (McGraw-Hill, New York, 1973).
23. Culver, C.G. *Am. Soc. Civ. Eng. Struct. Div. J.* **102**, 2269-2284 (1976).
24. Issen, L.A. *Single Family Residential Fire and Live Loads Survey* (NTIS PB 81-166647, 1980).
25. Larson, D.A. & Small, R.D. *Analysis of the Large Urban Fire Environment*, Part 2 (NTIS AD A125 400/2, 1982).
26. *Oil Market Report 2 May 1984* (International Energy Administration, Paris).

Is the "nuclear winter" real?

The discussion on the so-called "nuclear winter" is continued here, with further contributions to come in future issues of Nature.

THE article by Covey, Schneider and Thompson¹ breaks new ground in the atmospheric modelling of the possible effects of nuclear war, but despite that article and that of Turco *et al.*² two crucial questions remain: Will there really be a major cooling of the continental land masses (down to some -40°C)? And would such a cooling effect be maintained for months or seasons rather than days or weeks? Here I highlight outstanding uncertainties that need investigation.

Covey *et al.* start their model calculations by assuming a particulate layer in place in the atmosphere as a result of conflagrations of cities and forests following a nuclear exchange. This layer of smoke and soot is assumed to be more or less uniformly distributed over altitudes of 1 to 10 km; thick enough (optical depth = 3.0) to cut off the (shortwave) solar radiation from the Earth's surface but optically thin in the infrared (IR) (so that the surface can cool by radiating longwave radiation into space). In other words, they accept the physical model of Turco *et al.* and, not surprisingly, get a similar answer.

I have the following points to make about the certainty of these effects.

(1) The temperature change of the land surface (assuming small heat capacity) depends on a temporary imbalance between two very small fluxes: the inflow of solar radiation and the outflow of IR radiation, as well as the inflow and outflow of other energy fluxes. One cannot be sure even about the sign of the temperature change until all energy fluxes have been fully specified. Turco *et al.* and Covey *et al.* consider only a severe reduction of visible solar radiation, and explicitly assume that the particle layer would be optically thin in the IR, would not affect IR transmission so would not create a "greenhouse" effect.

(2) But the size and sign of the temperature change depend crucially on the ratio of the optical thicknesses (in the visible and the IR), and thus not only on the total amount but also on the size distribution of the particulates, especially on their detailed optical properties. The particle size distribution would be modified by agglomeration which is most important initially, when the smoke, soot and ash clouds are dense. For the same initial mass, the coalescence of smaller into larger particles will increase optical depth in the IR and decrease it in the visible.

(3) The conflagration would probably generate a variety of complex gaseous combustion products with absorption bands throughout the IR region.

(4) Water droplet clouds (located below the hypothetical particle layer) would trap

and re-emit IR radiation emanating from the land surface below.

(5) About the generation of water vapour there can be no doubt, although the phenomenon was not considered by Turco *et al.*

The combustion of dry mass must generate a corresponding mass of water vapour, since (CH_2) is oxidized to form H_2O . A total dry mass of 2.2×10^{13} kg (see ref. 3) must lead to $\sim 2.5 \times 10^{13}$ kg of water vapour. The amount of water vaporized by the nuclear explosion itself is small: the energy of a 1 Mtonne explosion ($\sim 4 \times 10^{15}$ J), if converted at the rate of 1 per cent, would lead to 1.6×10^7 kg of H_2O . Quite important, however, is the conversion of the energy released in the combustion: $(2.2 \times 10^{13}$ kg of dry mass) $(14 \times 10^6 \text{ J kg}^{-1}) \approx 3 \times 10^{20}$ J. A 25 per cent conversion would generate another 2.5×10^{13} kg of H_2O .

If the fire spread over an area of $\sim 10^6$ km², (ref. 3), it would produce from 2.5 to 5 g cm⁻² of water vapour (2.5 to 5 cm at STP), adding substantially to the normal water content of the atmosphere of about 2.5 g cm⁻². Such a large release of water vapour is possible, depending on the state of the forest, soil moisture, existence of streams and lakes, etc. In forest fires steam is produced copiously, as well as clouds, making it difficult to spot them from IR detectors on weather satellites^{4,5}. Eventually the water vapour will spread over a larger area, making it less effective as an IR absorber.

(6) Another neglected heat source, warming the Earth's surface, is the combustion energy of the very materials that contribute to the smoke and soot. The oxidation of wood releases $14 \times 10^6 \text{ J kg}^{-1}$, and 1 kg m⁻² daily corresponds to nearly 200 W m⁻² — the average amount of solar energy at a cloudless low-latitude location. A smouldering combustion of only 1 ounce of material per square foot per day would generate about 50 W m⁻², many times the minimum solar energy (8 W m^{-2}) calculated by Turco *et al.* for their baseline case. This would be sufficient to keep the surface temperature from falling too low.

Thus surface temperatures are unlikely to fall very low. They could even increase if particle size distribution, water clouds and combustion gases are such as to throttle the loss of heat radiation from the surface. Most probably, temperatures will decrease by a few degrees (rather than tens of degrees).

I now turn to a discussion of the lifetime of particulates in the atmosphere.

(i) The one-dimensional model of Turco *et al.* cannot capture mesoscale effects in the

atmosphere. Nor, unfortunately, can the global three-dimensional models of Covey *et al.*; their resolution is too coarse. Yet mesoscale effects are likely to determine the residence time of the smoke and soot particles. Experience with mesoscale models⁶ argues strongly against the existence of a stable layer in the atmosphere where the solar energy is absorbed. On the contrary, one would expect violent convective activity even if the layer were initially uniform, leading to cumulus-type formation, thunderstorms, rain squalls and accelerated cleansing of the atmosphere.

(ii) Further convective activity (plus rain showers and atmospheric cleansing) comes from the strong temperature gradient at the ocean-land boundary. Again, the three-dimensional model cannot capture these mesoscale effects, although the induced monsoonal circulation is yet another factor which limits large temperature excursions on the continents.

(iii) With the residence time limited by mesoscale effects, it is proper to ask if there is enough time to disperse the particulates uniformly through the global atmosphere, or at least the Northern Hemisphere, as explicitly assumed by Turco *et al.* and Covey *et al.* The latter suggest that the radiative heating of the particles may affect the atmospheric circulation so as to speed up their dispersion. This may well be the case; although in the absence of a fully interactive, self-consistent calculation, one cannot be sure how important the effect is.

Perhaps reasonably uniform dispersion can be achieved before rain-out cleanses the atmosphere. More probably, one will see patchy clouds which thin out rapidly — hardly a cataclysmic nuclear winter.

S. Fred Singer

1. Covey, C., Schneider, S.H. & Thompson, S.L. *Nature* **308**, 21-25 (1984).
2. Turco, R.P., Toon, O.B., Ackerman, T., Pollack, J.B. & Sagan, C. *Science* **222**, 1283-1292 (1983).
3. Crutzen, P.J. & Birks, J.W. *Ambio* **11**, 115-125 (1982).
4. Singer, S.F. *J. Forestry* **60**, 860-862 (1962).
5. Matson, M., Schneider, S.R., Aldridge, B. & Satchwell, B. *Fire Detection using the NOAA-series Satellites* NOAA/NESDIS Tech Rep. April 1984.
6. Pielke, R.A. *Mesoscale Meteorological Modeling: An Introductory Survey* (Academic, New York, 1984).

COVEY AND CO-WORKERS REPLY — Dr Singer's comments on the plausibility of "nuclear winter" deserve serious consideration, perhaps in more detail than limited space allows.

We emphasize again that the work reported in our article¹ was designed to examine how a three-dimensional climate model would respond to a hypothetical war-generated smoke cloud similar to that

*Department of Environmental Sciences University of Virginia, Charlottesville, Virginia 22903, USA

used by Turco *et al.*² in one-dimensional simulations. As we said, there are many major improvements to be made in the climate modelling alone (not to mention other components of the problem) before we have a satisfactory simulation of the climatic effects of nuclear war aerosols.

While we agree with Singer that two crucial questions are indeed the intensity and duration of surface cooling over land, the cooling need not be down to -40°C nor last for several weeks or more to produce devastating agricultural damage.

When Singer uses the phrase "physical model" of Turco *et al.*, we presume he means the column loading of smoke. The smoke loading is, as we state, taken from the scientifically "conservative" value used by the US National Academy of Sciences Committee on Atmospheric Effects of Nuclear Explosions as its baseline case, giving the absorption optical depth of smoke spread over 30° to 70°N as 3. Other estimates³ give an absorption optical depth nearly three times as great over the same area.

Contrary to Singer's assertion, we do not simply "get a similar answer" to that of ref. 2. Our results are different, but readily understandable from the known differences between one- and three-dimensional climate models. Due to ocean thermal influences our land-surface cooling is, on the average, roughly half that of Turco *et al.*, and land areas near coasts do not cool nearly as much as in their model. However, in our calculations, continental interiors may cool by 40°C or more in July (similar to the land values of Turco *et al.*) but our land surfaces can cool more quickly, to below freezing — even near coasts — in only a few days. This is a consequence of our resolution and of the separation of the surface and lower atmospheric temperature calculations, illustrating that adding detail to the calculations will not necessarily make the problem less severe.

(1) It is not clear what Singer is saying when he discusses the surface energy balance and the computation of surface temperature. It is sufficient to say that all surface energy fluxes are explicitly computed and balanced in our model. Any initial arguments concerning the sign of the temperature change at the surface resulting from the radiative effects of widespread persistent high-altitude smoke clouds have now been resolved in favour of cooling. The lack of significant compensating greenhouse warming from the IR emission of high-altitude nuclear smoke is borne out by detailed radiative transfer calculations, by Turco *et al.* and others, based on typical observed smoke aerosol properties.

This conclusion is quite robust. Extensive sensitivity tests⁴ show that substantial greenhouse compensation requires smoke loadings about two orders of magnitude greater than any so far postulated — but even in this extreme case, the surface temperature cannot exceed the effective radiating temperature of the

atmosphere, $T_E = (F/\sigma)^{1/4}$, where F is net downward solar flux at the top of the atmosphere and σ is the Stefan-Boltzmann constant. T_E is generally much less than normal surface temperature, which comes about because, by radiative/convective processes, substantial amounts of solar energy are absorbed at an altitude below that at which IR radiation is effectively emitted to space. Because the solar absorption coefficient for smoke aerosols is about an order of magnitude greater than the IR absorption coefficient⁴, and because nuclear smoke clouds would probably have their tops in the mid to upper troposphere, the effective altitude of solar absorption would be above the effective height of IR emission. This, in essence, is why the radiative/convective calculations yield little compensating greenhouse warming from nuclear smoke. (2) Singer mentions the importance of smoke particle coagulation in the early dense smoke plumes, which we did discuss. Solar absorption optical depths for widely-spread smoke might eventually be reduced by a factor of ~ 2 if initial smoke particles, of radius $0.1\ \mu\text{m}$ and smaller, coagulate to form $0.2\ \mu\text{m}$ to $0.3\ \mu\text{m}$ size particles. But for plausible amounts and distributions of smoke, the IR optical depth is not large enough for a significant ameliorating effect on surface temperature.

(3) Of course, "greenhouse" gases such as C_2H_4 , C_3H_6 and CO_2 will be generated by fires, but estimates of the amounts released are much too small for them to be capable of providing a substantial surface temperature compensating effect (J.T. Kiehl and J.W. Birks, personal communication).

(4) Singer is correct in noting that water clouds can have a large greenhouse effect. This effect was included in our model, though we cannot be certain that our — or anyone's — calculations can satisfactorily simulate what might actually happen to cloud cover, particularly near the surface.

(5) Singer's discussion of water vapour generated in burning areas is plausible, but must be reduced by two orders of magnitude when the water vapour is spread over a substantial fraction of a hemisphere. (We note that Singer does not mention one of the largest uncertainties in the "nuclear winter" scenario: the question of how much smoke from large fires will remain in the atmosphere after precipitation scavenging in cumulus clouds associated with the fires.)

(6) Singer's estimates of energy released in the fires are reasonable, but again the climatic impact is small when the effect is spread over a substantial fraction of a hemisphere. The total amount of energy released in fires from a large-scale nuclear war could be $\sim 3 \times 10^{20}\ \text{J}$ which, if released over two days and spread uniformly by winds over the entire land area between 30° and 70°N , would yield a power of $\sim 30\ \text{W m}^{-2}$. The heat available from smouldering fires would be of the order of $1\ \text{W m}^{-2}$, much less than the

absorbed solar energy of about 200 to 300 W m^{-2} , so that, on the average, the heat released from fires would do little to keep the land surface from cooling although areas near fires could receive significant heating as long as fires burned.

For all these reasons, we believe there is no sound basis for concluding that surface temperatures "are unlikely to fall very low". Certainly Singer has not offered quantitative grounds for preferring a decrease of "a few degrees" over "tens of degrees".

On the duration of the effects of nuclear winter we have the following comments:

(i) Mesoscale effects will be important in determining the removal of smoke from the atmosphere¹. Intuitively, one might expect enhanced moist convective activity around smoke clouds to increase the removal rate of the particles, but it is not yet possible explicitly to calculate the effects of dense smoke clouds on mesoscale circulations. In our model, the top of the smoke was the location of intense dry convective activity (as approximated by the model). Low relative humidity in the heated smoke prevented condensation.

(ii) Heat transport from ocean to land, as we stated, is indeed the reason why our predicted average temperature decrease is approximately half that of Turco *et al.*

(iii) Like Singer, we expressed concern about the global spread of smoke. But Singer does not pursue the importance of the height of the initial smoke plumes in determining the rate of removal of smoke from the atmosphere. If smoke is initially injected above most of the water vapour in the atmosphere (above 3 km) its lifetime will be much longer than if it is placed near the surface, where it can be more easily washed out.

The significance of patchy smoke clouds was discussed prominently in our article. It is not yet possible to make meaningful probability statements such as that it is "likely" that the patchy smoke clouds would thin out rapidly. Indeed, historical observations of some large fires⁵ show that substantial amounts of smoke can travel intercontinental distances in several days. But nothing has yet been found to undermine the plausibility of the notion that the smoke aerosols generated by a large-scale nuclear war could have lifetimes in the atmosphere of the order of weeks.

Starley L. Thompson*,
Stephen H. Schneider & Curt Covey

1. Covey, C., Schneider, S.H. & Thompson, S.L. *Nature* **308**, 21-25 (1984).

2. Turco, R.P., Toon, O.B., Ackerman, T., Pollack, J.B. & Sagan, C. *Science* **222**, 1283-1292 (1983).

3. Crutzen, P.J., Galbally, I.E. & Bruhl, C. *Climatic Change* (in the press).

4. Ramaswamy, V. & Kiehl, J.T. *J. geophys. Res.* (submitted).

5. Wexler, H. *Weatherwise* **3**, 129-142, December (1950).

*Authors addresses: National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA (S.L.T. and S.H.S.) and Rosentiel School of Marine and Atmospheric Science, University of Miami, Florida 33149, USA (C.C.)

Hunting for the missing mass

Neutrinos now seem to be poor candidates to supply the mass to close the Universe. Cosmologists should now deal with an observation suggesting that the problem of the missing mass is illusory.

THE cosmologists' hunt for what they call the missing mass is not just an hilarious entertainment mounted for the benefit of ordinary people. The underlying issue is whether the visible mass of the Universe, including that of the stars that constitute the galaxies, is substantially all there is. If so, the average density of the Universe is roughly $10^{31} \text{ g cm}^{-3}$, rather less than 10 per cent of the critical density needed to decelerate the expansion to zero — to bring it to a halt. Those who search for missing mass are therefore looking for some means of equipping the Universe with perhaps ten times as much matter as can now be seen. Whatever it is, it cannot by definition radiate, and must be in other ways inconspicuous.

The particular contribution of P. Hut and S.D.M. White (p.637, this issue) is their judicious examination of whether neutrinos (and antineutrinos) will fill the bill. The conclusion is that, with the constraints of particle physics on the one hand and of cosmology on the other, two of the three known species of neutrinos must carry mass and that the heavier of them, at least, must be radioactive. There are two ways of regarding this result. Optimists will take it as a prediction to be confirmed. Pessimists, likely to be the vast majority, will take it as a sign that the missing mass, whatever it is, does not consist of neutrinos.

But why take all this trouble? Why now further strain credulity by asking whether the missing mass may be made of axions, or photinos, whose very existence remains to be demonstrated? Would it not be simpler if cosmologists abandoned their general belief that there is missing mass, reconciling themselves instead to the notion that the Universe will go on expanding and the galaxies receding from each other indefinitely?

The trouble, as Dennis Sciama said in a Royal Society lecture last year, is that the problem of the missing mass "will not go away" (*Proc. R. Soc. A* **394**, 1-17; 1984). Hut and White say that the roots of the problem are in part "philosophical" — and do not by that imply that they have a prejudice about the kind of Universe in which they would prefer to live. Put simply, their point is that the near-coincidence between the density of the Universe and the critical density that would be needed to close it, and which differs merely by an order of magnitude, cannot be a chance coincidence.

Sciama points out that the coincidence is even more startling than it seems. Discrepancies between the density of the Universe and the critical density can only grow with the passage of time. So if the difference now is represented by a factor of ten, it must have been quite minuscule at the very early stages in the evolution of the Universe. So (simplifying a little), near-equality of the density and the critical density must be a feature of the system.

This provides *post hoc* support for the fashionable notion of the "inflationary universe", most simply regarded as evolution in two distinct phases, in the first of which the full menagerie of material particles allowed by some correct Grand Unified Theory are transformed into each other, creating matter at such a rate that the density stays critical. Affectionately, Sciama notes the similarity between this state of affairs and Hoyle's steady-state universe of more than thirty years ago. This picture has the advantage also of explaining why the microwave background radiation is isotropic to within one part in 1,000. And it makes the search for missing mass respectable.

There are also observations which have that effect, broadly speaking of two kinds. First, it is known from observations of spiral galaxies that individual stars appear to experience gravitational forces greater than can be accounted for by the masses of other visible stars as well as gas and dust clouds. Even the motion of the Solar System perpendicular to the plane of the Galaxy seems to be governed by a gravitational force twice as great as that inferred from the known distribution of massive objects. And the in-plane velocities of stars in other spiral galaxies do not decrease with distance from the centre as quickly as expected, supporting the idea that spiral galaxies have massive haloes. The most recent but perhaps most persuasive evidence that there is missing mass on a substantial scale comes from the behaviour of clusters of galaxies, which appear uniformly to be more tightly bound gravitationally than the sizes of their members would suggest. So the hunt has seemed a hunt for something real.

What, then, will the mass-hunters make of the results of a discordant investigation just published. (Tyson, J.A., Valdes, F., Jarvis, J.F., & Mills, A.P., *Astrophys. J. Lett.* **281**, 59; 1984)? The design of this gigantic computer exercise (at Bell Laboratories) is based on the expectation

that light from a distant galaxy should be affected by the gravitational field of galaxies standing closer along the line of sight, with the result that, in general, the apparent shape of a distant galaxy should be distorted by those lying in the foreground. Tyson and his colleagues have worked with the images of galaxies on a series of 35 plates exposed at the prime focus of the 4 m telescope at Kitt Peak, and containing altogether more than 200,000 images of galaxies, down to the 24th magnitude.

The procedure is to separate the galaxies into two groups, consisting of more and less distant objects, and then to select pairs of galaxies (separated by less than a minute of arc), of which there were nearly 28,000. The effect of foreground galaxies on the shapes of distant objects is assessed by the computation of a combination of the second moments about the centroid of an image. Obviously the interpretation of this data must boil down to looking for significantly systematic departures from circularity or ellipticity along the vector to the putatively deflecting star. Because many of the distant images will be only a few pixels across, only large numbers will ensure a meaningful result.

For the mass-hunters, the outcome is disconcerting. Distant galaxies in pairs separated by more than 10 seconds of arc are undistorted, but smaller angular separations bring misshapen images of the more distant galaxies. Tyson and his colleagues say that their estimates (of the mass of the nearer galaxy) would have to be in error by two or three standard deviations to be consistent with the larger estimates of mass produced by other methods. Their best estimate of the density of the Universe is a mere three per cent of the critical density (with a two standard deviation chance that the two densities are the same).

What is to be made of this unexpected challenge? First, more analysis (which Tyson *et al.* promise) is needed; obviously the technique itself is full of promise. Tyson and his colleagues, in explanation of the conflict with measurements suggesting there is lots of missing mass, say that, hitherto, people have concentrated on the brighter galaxies. Not all those who hunt for missing mass are dismayed by what has happened. One well-known practitioner is attracted by the possibility that the Tyson result is correct, relishing the idea that "it's telling us something about gravity we didn't know before". **John Maddox**

Cancer

Mutant *ras* proteins and cell transformation

from Rob Newbold

A SIGNIFICANT proportion of human tumour cell lines and chemically-induced rodent tumours contain activated *ras* genes capable of inducing morphological and tumorigenic transformation when introduced into preneoplastic immortal cell lines by DNA-mediated gene transfer. In all cases so far examined, activation depends on a point mutation in either codon 12 or codon 61 of the normal *ras* gene, resulting in a single amino acid substitution in the *ras* protein it encodes. The protein products of all three members of the *ras* gene family (designated Harvey-, Kirsten- and N-*ras* on the basis of homology with the two sarcoma virus oncogenes and the prototype SK-N-SH neuroblastoma oncogene respectively) are immunologically related and, having a molecular weight of approximately 21,000, are known as p21 proteins. Little is known about the function of normal p21 *ras* or about the way in which either of the specific amino acid substitutions generates proteins with such powerful transforming properties. What is certain is that the intracellular location of both normal and activated p21 *ras* is at the inner face of the plasma membrane. This fact has led to a general feeling that the proteins could play a major part in transducing signals from growth factors acting at the cell surface to the rest of the cell, and that this function might somehow be disturbed in the mutant transforming protein. A paper in this issue of *Nature* provides the first evidence of a biochemical difference between normal and activated *ras* proteins which could be directly relevant to the transforming ability of the latter¹.

Until now, the only known biochemical activity of p21 *ras* proteins was their capacity to bind a single guanine nucleotide (GTP or GDP) with high affinity. Although the site of nucleotide binding has not been unequivocally determined, those regions involved in the activation of p21 *ras* are implicated since they share sequence

similarity with the nucleotide-binding sites of analogous proteins^{2,3}. Furthermore, Harvey and Kirsten sarcoma virus p21s are known to phosphorylate their own threonine residue at position 59, which is close to the activation site at position 61. Clearly, though, the guanine nucleotide-binding activity of the p21 proteins is not directly involved in activation as neither the extent nor the specificity of binding seems to be altered by mutations that confer transforming activity⁴.

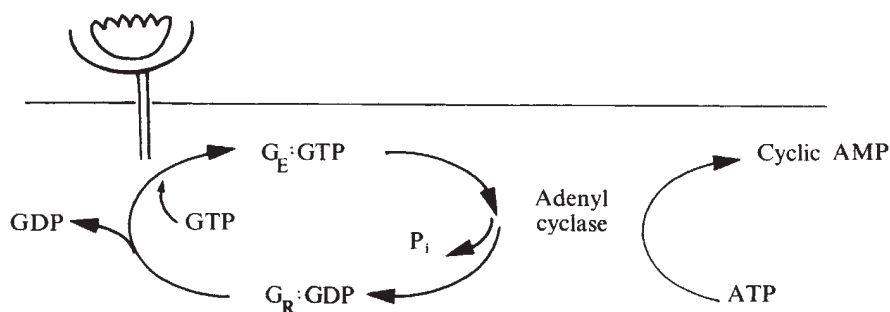
Arthur Levinson and his colleagues at Genentech have therefore undertaken a more rigorous characterization of normal and activated p21 *ras* proteins with the aim of uncovering any vital differences that could account for the latter's transforming potential; their results¹ are reported on page 644. In order to produce enough material for their study, the Genentech group constructed plasmids containing wild-type or mutant (codon 12) human Ha-*ras* cDNAs under the control of the *trp* promoter such that the genes were efficiently expressed in *Escherichia coli*. Inspired by the fact that other known purine nucleotide-binding proteins possess the ability to catalyse the removal of the terminal phosphate group from nucleotide phosphates, McGrath *et al.* have looked for similar activity in p21 *ras* by examining the capacity of *E. coli* extracts containing p21 to hydrolyse GTP to GDP. Their exciting discovery is that while normal p21 *ras* has a GTPase activity, this function is grossly impaired in the oncogenic protein. This finding is confirmed in another paper⁵ shortly to appear in *Nature*. Thus, a substantial difference between wild-type and activated p21 *ras* has finally emerged, and it prompts speculation as to how the proteins function in normal and transformed cells.

Two other recent studies provide further insight into *ras* proto-oncogene activation. Fasano *et al.*⁶ have looked for novel activating mutations in p21 protein. This

was done by performing random *in vitro* mutagenesis of a cloned wild-type Ha-*ras* gene using sodium bisulphite, and testing the mutants for transforming activity in NIH 3T3 cells. DNA sequencing of the transforming mutants revealed three new targets for activation (13, 59 and 63) flanking codons 12 and 61. Interestingly, this means that the viral *ras* proteins, v-Ha-*ras* and v-Ki-*ras*, have two activating missense mutations as they have a threonine substitution at residue 59 in addition to a substitution at position 12. But Fasano *et al.* go on to show, using chimaeric plasmids, that the second mutation does not enhance transforming efficiency. What is not clear is why *ras* oncogenes with mutations other than in codons 12 and 61 have not yet been detected in tumour cells, although, as the authors point out, all mutants capable of altering the morphology of NIH 3T3 cells may not be equally effective at inducing malignant transformation.

The concept of qualitative differences in the transforming potential of mutant *ras* proteins finds further support in a second paper from Levinson's group⁷, shortly to be published in *Nature*. In this study, human Ha-*ras* mutants were constructed which specified all 20 possible amino acids at position 12 of p21. Seeburg *et al.* find that with the exception of proline, every substitution for glycine generates a protein capable of morphologically transforming Rat 1 cells. The degree of transformation induced by each of the mutants is highly variable, however, leading the authors to propose that in using this type of transformation assay we may so far have underestimated the proportion of tumours containing activated *ras* genes. The next step will be to see which mutations affect the GTPase activity of p21 *ras*.

What is the significance of the new findings about p21 *ras* function and activation? The key observation is the similarity between p21 *ras* and the guanine nucleotide-binding proteins (termed G- or N-proteins) which are responsible for stimulation and inhibition of adenylate cyclase following the attachment of polypeptide hormones to specific cell-surface receptors (ref. 8 and see the figure). In addition to possessing a high-affinity guanine nucleotide-binding site, the G-proteins share sequence similarities with p21 *ras* and also have GTPase activity which regulates their function. In the case of the stimulatory arm of the adenylate cyclase system, binding of the relevant hormone to its receptor promotes conversion of the inactive GDP form of the G-protein into the GTP state which in turn activates the cyclase. Thereafter, increased production of cyclic AMP (the second messenger) stimulates a cyclic AMP-dependent protein kinase which carries out specific regulatory phosphorylations that determine the response of the cell. The G-protein signal is cancelled by slow hydrolysis of GTP to GDP, mediated by



Signal transmission from cell-surface receptor to adenylate cyclase through G-protein. The activated (effector) GTP-bound form of G-protein, $G_E:GTP$, which activates adenyl cyclase, is converted to the inactive form, $G_R:GDP$, by GTPase. *ras* protein is hypothesized to act like G-protein.

the GTPase. G-proteins consist of three subunits and it is thought that GTP-binding promotes their reversible dissociation, releasing the large α -subunit which acts as the effector.

The new papers contain confirmation of the earlier observation⁹ that amino acid substitutions in p21 *ras* transforming proteins result in altered electrophoretic mobility on SDS-PAGE gels, suggesting gross conformational changes. Furthermore, it has been proposed on the basis of computer modelling¹⁰ that these mutant forms of p21 *ras* would be more conformationally restricted than the wild-type protein. If p21 *ras* does act as a signal transducer in a similar fashion to the G-proteins, then it is easy to see how loss of GTPase activity and/or a more rigid conformation could result in a permanent signal in the absence of the ligand. In fact a remarkably analogous situation occurs in the adenylate cyclase system when cells are exposed to cholera toxin. In this case, part of the toxin ADP-ribosylates a specific site on the G-protein, inhibiting its GTPase activity and thereby fixing it in the GTP-form. The abnormally high levels of intracellular cyclic AMP which result are responsible for the devastating effects that cholera toxin has on active transport systems in the intestinal mucosa. By the same mechanism, cholera toxin also exerts a strong proliferative stimulus on certain mammalian cells in culture.

Although speculation about specific membrane-bound effector systems involving *ras* proteins is probably premature, a functional link does appear to exist between p21 *ras* and receptor systems stimulated by epidermal growth factor (EGF) and insulin, as both growth factors have been found to enhance guanine nucleotide binding by *ras* proteins in isolated membranes¹¹. It should be noted, though, that the EGF receptor probably does not rely on G-like proteins for signal transduction (at least for activation of its tyrosine-specific protein kinase) since binding, transducing and catalytic activities are contained in a single polypeptide. Nonetheless, it is becoming increasingly likely that transforming *ras* proteins will turn out to function by subverting growth factor controls on cell proliferation. However the crucial question concerning when and to what extent such relaxation of cell-cycle controls contributes to the complex multistage process of carcinogenesis remains unresolved. □

1. McGrath, J.P., Capon, D.J., Goeddel, D.V. & Levinson, A.D. *Nature* **310**, 644 (1984).

2. Gay, N. & Walker, J. *Nature* **301**, 262 (1983).

3. Wierenga, R. & Hol, W. *Nature* **302**, 842 (1983).

4. Finkel, T. *et al.* *Cell* **37**, 151 (1984).

5. Sweet, R. *et al.* *Nature* (in the press).

6. Fasano, O. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 4008 (1984).

7. Seeburg, P.H. *et al.* *Nature* (in the press).

8. Gilman, A. *Cell* **36**, 577 (1984).

9. Der, C. & Cooper, G. *Cell* **32**, 201 (1983).

10. Pincus, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 5253 (1983).

11. Kamata, T. & Feramisco, J. *Nature* **310**, 147 (1984).

Rob Newbold is at the Institute of Cancer Research, Fulham Road, London SW3 6JB.

Cell biology

Multiple controls for secretion?

from Peter F. Baker

EVERY student knows that secretion requires calcium, especially when the substances to be secreted — neurotransmitters, hormones or enzymes — are packaged within intracellular vesicles, the contents of which are released by exocytosis. So it may seem surprising to find a trickle, rapidly becoming a spate, of papers, including one by di Virgilio, Lew and Pozzan on page 691 of this issue of *Nature*, questioning the primacy of calcium in secretory control¹⁻⁷. These have been triggered by two kinds of experiment: direct measurement using the fluorescent dye Quin 2 of intracellular free Ca during secretion — something that had hitherto been possible only in a few large cells — and the discovery that exposure of cells to certain phorbol esters can evoke exocytosis apparently with little or no change in intracellular free calcium.

Quin 2 has revealed a variety of calcium responses to secretory stimuli. One side of the coin is represented by adrenal medullary cells⁸ and mast cells⁶ which show a large rise in free Ca in response to secretagogues such as aequorin revealed a large change in ionized Ca following electrical stimulation of the squid giant axon synapse. The other by the platelet which exhibits a rise in Ca to some stimuli but little or no change in response to other equally effective secretagogues⁴. Furthermore, treatment of platelets and a variety of other secretory cells with the phorbol ester, 12-*O*-tetradecanoylphorbol 13-acetate (TPA), tends to promote exocytosis often without detectable change in free calcium. Almost without exception, after pretreatment with phorbol ester, subsequent exposure to low concentrations of a calcium ionophore reveals that the secretory system has been sensitized to low concentrations of calcium.

The experiments with phorbol esters offer a clue to mechanism. Phorbol esters have been shown to substitute for diacylglycerol in the activation of the enzyme protein kinase C⁹ and this enzyme is the only phorbol receptor so far identified¹⁰. It follows that the enzyme may be involved in the control of exocytosis, a suggestion that is supported by the close quantitative parallel between those phorbol esters that activate protein kinase C and the activation of secretion. If phorbol esters can evoke exocytosis, presumably the endogenous compound diacylglycerol for which they substitute can do so also and secretagogues which effect exocytosis with minimal change in ionized Ca may act to promote phospholipid breakdown, thereby raising the local concentration of diacylglycerol near the C-kinase.

The key question from the viewpoint of

secretory control is whether phorbol ester-induced secretion reflects activation of a parallel pathway completely independent of that used by calcium. Although some authors have made such assertions, their claims may be premature. There seem to be at least three very distinct possibilities: (1) that Ca and phorbol esters act to promote exocytosis by two quite separate and independent systems; (2) that secretion requires Ca and the sensitivity of this reaction can be modified by protein kinase C; and (3) that the actions of both Ca and phorbol esters are channelled through protein kinase C. Only the first possibility provides for two truly separate pathways. The other two utilize protein kinase C either as a modulator of a basic Ca-dependent process or as a final common pathway for secretory control.

Assessment of these possibilities relies heavily on the properties of protein kinase C and here much still remains to be done. For instance, can characteristics determined in the chaos of the test tube, where full activation requires Ca, diacylglycerol and a phospholipid, be extrapolated to the highly structured intracellular environment? Even the physiological substrates of the kinase are largely unknown, making *in vitro* studies, at best, those of a model system. Does the enzyme activity disappear completely in the presence of a strong Ca-chelating agent? Or can it still phosphorylate some substrates under such conditions? Only more experiments will tell.

Because of the uncertainties surrounding the *in situ* biochemistry of C-kinase, another approach to the analysis of secretory control seems desirable. A rather direct approach is possible with electrically permeabilized secretory cells¹¹. In these preparations the intracellular free Ca can be controlled and the dependence of secretion on intracellular free Ca determined directly. Activation is normally half-maximal at about 1 μ M Ca. In all the cell types so far examined, adrenal medullary cells¹², platelets¹³ and exocrine pancreas¹⁴, exposure to phorbol esters sensitizes secretion to Ca by promoting a leftward shift — by up to one log unit — of the intracellular Ca activation curve; but in all cases phorbol esters are ineffective at very low buffered levels of calcium. The leftward shift in permeable platelets seems quantitatively large enough to account for activation of secretion by phorbol esters in intact cells where Ca does not change. These findings are more compatible with possibilities (2) and (3) than with possibility (1).

Nevertheless, not all the observed effects of phorbol esters on secretion may be explicable in this way. Only experiments of the kind described above on permeabilized

cells can permit a clear answer; but there are already examples where the leftward shift will have to be very large to accommodate the observed findings. For instance, di Virgilio, Lew and Pozzan¹ report that phorbol esters can promote secretion from neutrophils at an ionized Ca concentration of 10 nM. This would require a bigger shift in sensitivity than has so far been observed in 'leaky' cells; only determination of the properties of C-kinase in neutrophils can substantiate or refute the authors claim for a completely Ca-independent mechanism.

Although this particular case must be considered non-proven, secretion — like contraction — seems likely to involve multiple controls. Apart from the release of stored products, membrane fusion plays a central role in the control of membrane turnover and the nature of the signals involved in directing membrane traffic are obscure. Even within the limited area of exocytotic secretion there are tissues that appear to secrete in response to a fall in ionized calcium. In one of these, the parathyroid, such changes have been very clearly documented by Quin 2 (ref. 15) and there is strong circumstantial evidence pointing

to something similar in the control of renin release¹⁶. Whether such findings necessitate something quite different in the way of secretory control or can be accommodated within more conventional systems — perhaps by inclusion of a Ca-dependent inhibitory control — are matters for future experiments to resolve. □

1. Di Virgilio, F.D., Lew, D.P. & Pozzan, T. *Nature* **310**, 691 (1984).
2. Gunther, G.R. *J. biol. Chem.* **256**, 12040 (1981).
3. Yamanishi, J. *et al. Biochem. biophys. Res. Commun.* **112**, 778 (1983).
4. Rink, T.J., Sanchez, A. & Hallam, T.J. *Nature* **305**, 317 (1983).
5. Wooten, M.W. & Wrenn, R.W. *FEBS Lett.* **171**, 183 (1984).
6. White, J.R. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3978 (1984).
7. Pozzan, T. *et al. J. Cell Biol.* (in the press).
8. Knight, D.E. & Kesteven, N.T. *Proc. R. Soc. B* **218**, 177 (1983).
9. Castagna, M. *et al. J. biol. Chem.* **257**, 7847 (1982).
10. Nield, J.E., Kuhn, L.J. & Vandenberg, G.R. *Proc. natn. Acad. Sci. U.S.A.* **80**, 36 (1983).
11. Baker, P.F. & Knight, D.E. *Nature* **276**, 620 (1978).
12. Knight, D.E. & Baker, P.F. *FEBS Lett.* **160**, 98 (1983).
13. Knight, D.E. & Scrutton, M.C. *Nature* **309**, 66 (1984).
14. Knight, D.E. & Koh, E. *Cell Calcium* (in the press).
15. Shoback, D.M. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3113 (1984).
16. Keeton, T.K. & Campbell, W.B. *Pharmac. Rev.* **32**, 81 (1981).

Peter F. Baker is Head of the Department of Physiology, King's College, London WC2R 2LS.

some investigators believe that these problems may be so serious that only the secular changes recorded on a few heavily travelled ship routes may be genuine (Ramage, C.S. *J. Climate appl. Meteor.* **23**; Feb 1984, and Barnett, T.P. *Mon. Weath. Rev.* **112** (2); Feb 1984). If real, the change near 1900 suggests that a sharp adjustment of the global system occurred, one which should be reflected in other independently observed geophysical fields such as surface wind and surface pressure. The combined record of all the surface fields can be viewed with greater confidence than one field alone. The pressure field must relate to the wind field, and in regions of strong sea-surface temperature gradients the wind field relates to sea-surface temperature anomalies.

Indeed, the record of the surface wind field does reveal adjustments of the global ocean — atmosphere system around 1900 and in the 1930s. Changes in the wind field during the 1900 adjustment of the global circulation have been partially described (Fletcher, J.O. & Fu, C. *Proc. 7th Climate Diagnostics Workshop*, US Dept of Commerce, NOAA; Oct 1983). They include a decrease in the vigour of the surface wind field over most of the ocean domain, with the most abrupt change in the Asian sector of the Southern Hemisphere westerlies, and an abrupt adjustment of the seasonal position of the tropical convergence in the far western tropical Pacific. Indirect evidence for an adjustment of the large-scale atmospheric circulation in the Indian Ocean around 1900 is also found in the African record, as a sharp decrease in rainfall feeding the Nile River (Riehl, H. & Meiten, J. *Science* **206**, 1178; 1979), and on the Antarctic side, as a sudden warming at Mizuho station, signalled by oxygen isotope ratios in ice cores (Kato, K. in *Ice Caring Project at Mizuho Station*, ed. Kusunaki, K., 165, National Institute of Polar Research, Tokyo; 1978). The implications of the sea-surface temperature record should not, therefore, be dismissed.

The global marine surface data set contains the most detailed record we will ever have of the dynamics of the global climate system over the last century and more. During this period several sharp adjustments seem to have occurred. What remains to be done now is to delineate more fully the spatial and temporal characteristics of these adjustments and to glean from them clues to the nature and causes of global climate change. With the aid of recent work on sea-surface fields, radiosonde coverage since about 1950, satellite observations since about 1970 and, more recently, three-dimensional modelling of the behaviour of the atmosphere, the implications of the long historical record should come clear. □

Joseph O. Fletcher is Assistant Administrator for Research in the National Oceanic and Atmospheric Administration, Washington, DC 20852.

Climatology

Clues from sea-surface records

from Joseph O. Fletcher

THE paper by Folland, Parker and Kates in this issue of *Nature* (page 670) gives important new insights into the behaviour of the global climate system. Their records show that globally averaged sea-surface temperature not only underwent a large fluctuation over the last century but that the large changes — cooling about 1900 and warming in the 1930s — were sudden rather than gradual.

The rapid change found in historic sea-surface records is good news for investigators of climate variability because the existence of large and abrupt 'climate signals' simplifies the strategy for investigation. Instead of groping in statistics to extract signal from noise, researchers can focus on the large signals and ask questions aimed at revealing the dynamics of climate variability on that time scale. Where does the signal appear first? How does it spread? What is the sequence of its evolution? Is the signal apparent in other geophysical fields such as surface wind and surface pressure? If so, do the changes in the various fields relate to one another in physically consistent ways? What do the trends tell us about possible causes and influential mechanisms?

The sea-surface record offers the best opportunity for delineating the dynamical behaviour of the global system over the last century — and it is a remarkably good record. Thanks to international agreements and standardized procedures

adopted in 1853 (Whipple, A. *Smithsonian*, 171; March 1984), about 100 million observations have been taken by ships over the last 130 years. The quantities recorded include sea-surface temperature, wind direction and speed, atmospheric pressure and temperature, sea state and cloudiness. Comparisons between the fields offer additional dimensions for judging the validity of signals reflected in individual fields.

The data are unevenly distributed in time and space with large gaps in the early years. About half of the total observations were made in the Atlantic, which is the most fully surveyed ocean area. The Pacific is the least covered, with especially large gaps in the data for early decades. By contrast, the record is excellent in the belt of westerlies in the Indian Ocean and South Atlantic. Because the sailing route, during the last century, from Europe to the Indies and the Orient was typically south from Europe to the Southern Hemisphere westerly belt near Argentina and then east, coverage in this region was as good in the 1850s to 1870s as a century later. As a result, changes in sea-surface temperature and global surface wind field are well delineated since 1854.

The most striking feature of the global ocean-surface temperature record is the sharp cooling shortly after 1900. Folland *et al.* point out many possible sources of unreliability of the data, however, and

Mineralogy

The riches of silica revealed

from Geoffrey D. Price

DESPITE its chemical simplicity, silicon dioxide (silica) exhibits a complex array of structural and material behaviour. In geology, SiO_2 has been widely studied because, as quartz, it is one of the most abundant minerals in the Earth's crust. Silica is equally important industrially, being used in the form of ceramics, gemstones, molecular sieves, oscillators, piezoelectric crystals and glass. Furthermore, recent reports on various forms of SiO_2 show that its study can still provide new insights into other diverse fields of research, ranging from the theory of solid-state transformations to the extinction of the dinosaurs.

Silica can adopt many crystalline structures, or polymorphs, the more important of which include α - and β -quartz, tridymite, cristobalite, coesite and stishovite. Each of these phases has a well defined field of thermodynamic stability, α -quartz being stable under ambient conditions, while β -quartz, tridymite and cristobalite become stable with increasing temperature. Coesite and stishovite are stable only at high pressures. All these polymorphs have quite distinct structures, although the phases stable at lower pressures are all built from SiO_4 tetrahedra, which are linked at each of their corners with another tetrahedron, thus forming a three-dimensional framework. Only in stishovite is a more dense packing of atoms observed, with silicon in 6-fold coordinated sites. The properties of a hypothetical SiO_2 polymorph, having a density even greater than that of stishovite and silicon in 8-fold coordinated sites, have recently been investigated theoretically by Carlson *et al.* (*Geophys. Res. Lett.* **11**, 617; 1984). It has been suggested that such a polymorph may occur at depth within the Earth's mantle or core, although Carlson *et al.*'s calculations do not support this.

The transformations between silica polymorphs, which involve the breaking of Si-O bonds, are invariably sluggish because of the relatively large amount of energy required to reconstruct the silica framework. The slow transformation kinetics allow the development of a large number of metastable phases such as glasses, the various forms of opal, crystalline polymorphs such as melanophlogite and silicalite, a zeolite-related hydrophobic molecular sieve (see Price, G.D. *et al.* *Nature* **292**, 818; 1981). Furthermore, once formed, high-temperature or high-pressure polymorphs can exist metastably outside their stability fields, even on the geological time scale. At lower temperatures, the metastable phases frequently undergo displacive phase transformations, in which their structures

become slightly distorted, thus adding to the complexity of silica phase relations.

It is a displacive phase transformation that relates the thermodynamically stable α - and β -quartz polymorphs. Until quite recently, this transformation was thought to be well understood, and was often quoted in text-books as a classic example of its type. Detailed diffraction and calorimetric studies, however, have shown that the transformation is far from simple. For example, Zeyen *et al.* discovered that a thermodynamically stable incommensurate SiO_2 phase exists over a temperature interval of a few degrees between the fields of stability of α - and β -quartz (*Physica B120*, 280; 1983). Quartz is thus one of the increasing number of insulators which exhibit incommensurate behaviour, a phenomenon that is providing a great theoretical challenge and which is at the forefront of research in solid-state physics (see Heine, V. & McConnell, J.D.C. *J. Phys. C17*, 1199; 1984).

Geologically, the persistence of SiO_2 polymorphs in a metastable state is particularly valuable because these phases provide a partial record of the pressure and temperature history of the rock in which they occur. In a spectacular recent X-ray and electron microscope study of quartz grains from a claystone band occurring at the sharply defined Cretaceous-Tertiary stratigraphic boundary, Bohor *et al.* (*Science* **224**, 867; 1984) found that this quartz showed features characteristic of shock effects, and was associated with traces of the high-density stishovite polymorph. They conclude that the claystone layer was produced by fall-out from the impact of a large extraterrestrial body on a Cretaceous landmass. If similar features were to be found around the world, this would constitute strong evidence that the end of the Cretaceous was marked by a number of extensive and catastrophic meteoritic impacts, which have previously been suggested as the cause of the inferred mass extinction of a variety of flora and fauna at this time (see *Science* **223**, 1174; 1983 and *Nature* **308**, 685; 1984).

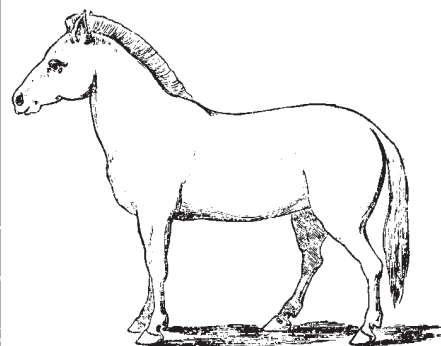
Observations of SiO_2 in the form of coesite may also give clues to the role and behaviour of continental crustal material in geodynamic and plate tectonic processes. Work by Chopin (*Contr. Miner. Petrol.* **86**, 107; 1984) on some extraordinary rocks from the western Alps has revealed the existence of coesite inclusions within 25-mm-diameter pyrope garnets. The coesite appears not to have inverted to a lower-density SiO_2 polymorph because of confining pressure applied by the surrounding garnet. A

similar discovery of coesite inclusions in clinopyroxenes in a dolomite-eclogite from Grytting in Norway lead Smith (this issue of *Nature*, page 641) to suggest that the eclogite had been subjected to pressures of ~ 30 kbar. This is an extraordinarily high pressure for crustal rocks to be exposed to, and suggests that the rocks have either been buried or subducted to depths of ~ 100 km, or have been subjected to significant overpressuring during deformation at a more shallow depth.

The importance of studies on SiO_2 phases is underlined by a recent edition of the *Journal of Geophysical Research* (**89B**(6); 1984) which contains nine papers dedicated solely to the study of the mechanical properties of quartz. Because of its ubiquitous occurrence, quartz has traditionally been treated as an analogue of continental crustal material. Major uncertainties about its mechanical behaviour still exist however. Quartz is one of the strongest common natural materials, yet, paradoxically, there is abundant microscopic evidence in quartz-bearing rocks that quartz has flowed extensively in nature and was weaker than other minerals associated with it. It has long been recognized that quartz is greatly weakened in the presence of water, a phenomenon known as hydrolytic weakening and this is thought to be responsible for the apparent weakness of natural quartz. The mechanism by which hydrolytic weakening occurs, however, is still the subject of a considerable research effort (see Blacic R.D. and Christie J.M., *J. Geophys. Res.* **89B**(6) 4223; 1984). □

Geoffrey D. Price is a Royal Society Research Fellow at the Department of Geology, University College, London WC1E 6BT.

100 years ago Przevalsky's Wild Horse



GREAT interest is attached to the question of the origin of our domestic animals, and especially to that of the horse — which is generally supposed not now to exist in an aboriginally wild state. Every fact bearing upon this subject is of importance, and the discovery by the great Russian traveller Przevalsky, of a new wild horse, more nearly allied to the domestic horse than any previously known species, is certainly well worthy of attention.

From *Nature* **30**, 391, 21 August 1884.

Evolutionary biology

Evolution of stable exploiter-victim systems

from Jared M. Diamond

PARASITES, predators and pathogens make their living by killing or weakening other organisms. Whenever such exploiter species expand their range, they encounter naive potential victims which lack behavioural defences. The frequent result is decimation or extermination of the victimized populations, as illustrated by the effects of smallpox on many human populations, of European chestnut blight on American Chestnut and of introduced cats and rats on dozens of island bird species. Why do exploiters not follow their victims into extinction? Why does the world instead abound with predators and their prey, with parasites or pathogens and their hosts? Four recent studies of parasitic vertebrates illuminate this evolutionary paradox.

The first study involves an old and stable parasite-host dyad: the Shining Cuckoo (*Chrysococcyx l. lucidus*) and Grey Warbler (*Gerygone i. igata*) of New Zealand^{1,2}. The cuckoo lays its eggs in the nest of the warbler but of no other New Zealand bird. It removes one warbler egg and substitutes its own, and the baby cuckoo, on hatching, proceeds to expel all remaining warbler eggs and nestlings from the nest. The warbler has no behavioural defences against the cuckoo; it accepts and rears the cuckoo nestling. Why, then, is the warbler still abundant and the cuckoo common?

Each year the warbler lays two clutches of four eggs each, the first in mid-September, the second in mid-November. The cuckoo migrates to remote Pacific islands for the austral winter and does not return to New Zealand until the end of September. The warbler's first clutch therefore escapes the cuckoo, and 54 per cent of the eggs yield fledged young. For the unparasitized second clutches the fledging rate is 33 per cent but cuckoos parasitize 55 per cent of the remainder and reduce the fledging rate to 2 per cent. Hence a pair of warblers rears about $(4 \times 0.54) + 4[(0.55 \times 0.02) + (0.45 \times 0.33)] = 2.8$ fledged young per year, many more than is needed to maintain the population of this sedentary long-lived species. A pair of cuckoos produce about 16 eggs per year, and the fledging rate is a healthy 52 per cent, yielding eight fledged young to make good the losses from the 8,000-km over-water return flight to the wintering grounds. Thus, both the cuckoo and its host thrive.

Let us now turn to the young unstable host-parasite systems involving the Brown-headed Cowbird (*Molothrus ater*) of North America³⁻⁵. This nest parasite,

which victimizes over 100 species of songbirds, probably evolved in the Great Plains and followed the vast herds of buffalo, eating insects stirred up by the herds. With the arrival of Europeans and near-extermination of buffalo, cowbirds switched to following horses and cattle and spread to other areas of North America, encountering songbirds naive to cowbirds. The result has been the decimation of some of the cowbird's new victims, including Kirtland's Warbler (*Dendroica kirtlandii*) in Michigan^{6,7} and the Bell's Vireo population of California (*Vireo belli pusillus*)^{8,9}.

Kirtland's Warbler is confined to young jack-pine stands of lower Michigan. The first recorded parasitization of a Kirtland's Warbler nest by cowbirds was in 1908. Thereafter, the proportion of parasitized nests rose rapidly to 25 per cent by 1931-50 and to over 70 per cent by 1966-71. The warbler is an ideal host: its ground nests are easy for cowbirds to find, the warblers accept cowbird eggs and tend them diligently, and warbler eggs hatch later than cowbird eggs and yield smaller nestlings unable to compete with the baby cowbird. Even in the absence of nest parasitization, disturbance by cowbirds reduces the number of warbler eggs laid from five to three. When the 1971 breeding census revealed a 60 per cent drop in the warbler population to 201 pairs, the US Fish and Wildlife Service began a massive programme to kill cowbirds in warbler habitats and the decline of the warbler population was halted.

The case of the Californian Bell's Vireo is similar^{8,9} but has one added twist^{10,11}. In 1908, the same year that Kirtland's Warbler was first victimized in Michigan, the first parasitized vireo nest was discovered. Having been abundant, the vireo was soon suffering from the highest rate of parasitism of any Californian songbird: 16 per cent of nests in 1908-17, 30 per cent in 1918-27 and 58 per cent from 1928 to present. Production of fledged Bell's Vireos dropped to 0.6 per nest and the species disappeared from most of California. Cowbirds have remained abundant in California by victimizing many other songbird species. Unlike the Kirtland's Warbler, the Bell's Vireo is not restricted to one habitat but has populations in midwestern North America, in areas long occupied by cowbirds. These cowbird-wise vireos suffer much lower rates of parasitism, produce young at two to three times the rate of the Californian population and are actually increasing their range^{10,11}.

The contrast between naive and experienced host populations is illustrated further by the effects of lampreys on trout

populations. Land-locked populations of the sea lamprey *Petromyzon marinus* have long coexisted in the lakes of upper New York State (Lake Ontario and the Finger Lakes) with the Lake Trout *Salvelinus namaycush*. While these experienced trout avoid lamprey attack and remain common, naive Brown Trout (*Salmo trutta*) introduced by fishermen are destroyed by lampreys. Originally, lampreys were prevented from reaching the upper Great Lakes by Niagara Falls, but construction of the Welland Canal allowed them to bypass the falls and reach Lake Erie in 1921 and Lake Michigan in 1936. Within a short time lampreys virtually exterminated the naive upper-lake population of Lake Trout as well as two other fish species, the ciscos *Coregonus johanna* and *C. nigripinnis*.

What light do these examples shed on exploiter-victim coexistence? In each case there was decimation of victims naive to their exploiters. Subsequently, one of four things may happen. (1) The exploiter may follow its victims into extinction, as illustrated by the cats that destroyed most of the birds of Mangere Island and then died out themselves¹². (2) The predator may survive by switching to victims that are pre-adapted to withstand the impact. For instance, the Grey Warbler does not have any behavioural defences against Shining Cuckoos, but nevertheless remains abundant because its first clutch is underway before cuckoos arrive. (3) Some victims survive long enough to evolve behavioural defences: trout learn to avoid lamprey attack, while songbirds learn to expel cuckoo eggs or abandon a parasitized nest. (4) The exploiter itself may evolve in such a way as to reduce its impact, as illustrated in host-pathogen systems by the evolution of lowered pathogen virulence as well as of host immunity.

The cowbird's recent expansion in North America represents a natural experiment offering rare opportunities to the evolutionary biologist. Cowbirds have been common in midwestern, northeastern, far western and southeastern North America for varying lengths of time and in each area there are dozens of songbird species that constitute actual or potential cowbird victims. Comparisons of the behaviour of cowbirds and their victims in these four areas may explain how stable exploiter-victim systems evolve. □

- Gill, B.J. *Ibis* 125, 40 (1983).
- Gill, B.J. *Notornis* 30, 137 (1983).
- Friedmann, H. *The Cowbirds* (Thomas, Springfield, 1929).
- Mayfield, H.F. *Am. Birds* 31, 107 (1977).
- Verner, J. & Ritter, L.V. *Auk* 100, 355 (1983).
- Mayfield, H. *The Kirtland's Warbler* (Cranbrook Inst. Sci., Bloomfield Hills, Michigan, 1960).
- Walkinshaw, L.H. *Kirtland's Warbler* (Cranbrook Inst. Sci., Bloomfield Hills, Michigan, 1983).
- Goldwasser, S. *Calif. Dept Fish & Game Rep., project W54-R10*, Job IV-1.5.1 (1978).
- Goldwasser, S., Gaines, D. & Wilbur, S.R. *Am. Birds* 34, 742 (1980).
- Nolan, V. *Condor* 62, 225 (1960).
- Overmire, P.G. *Condor* 64, 75 (1962).
- Diamond, J.M. *Nature* 309, 308 (1984).

Jared M. Diamond is Professor of Physiology at University of California School of Medicine, Los Angeles, California 90024.

Photosynthesis

Novel carbon supply on land

from Peter D. Moore

AN adequate supply of carbon is a fundamental need of all autotrophic organisms, and not surprisingly, an intriguing range of techniques have evolved to satisfy it. A number of aquatic plants have been reported to use a carbon assimilation pathway of the type normally associated with desert succulents, named crassulacean acid metabolism (CAM; see Moore, P. D. *Nature* **304**, 310; 1983), which permits accumulation of dissolved CO₂ and bicarbonate ions during the night, when there is a better supply than in daylight. The CO₂ supply in the dark ocean waters is tapped by chemolithotrophic microbes which obtain energy from the organic rain of dead material sinking from the euphotic zone (Karl, D. M., Knauer, G. A., Martin, J. H. & Ward, B. B. *Nature* **308**, 54; 1984). Terrestrial plants obtain their carbon from the atmosphere as gaseous CO₂, but often it is in such short supply that it limits the rate of photosynthesis. When the light supply is adequate, the rate of photosynthesis increases linearly with increasing CO₂ from 0.03 to about 0.1 parts per 10⁶ (Hall, D. O. & Rao, K. K. *Photosynthesis*, Edward Arnold, London). Obviously, therefore, it would be a selective advantage if a plant could locate alternative supplies of CO₂. What about the abundance of CO₂ released in the soil by decomposer and root respiration; would it not be advantageous for plants to supplement their atmospheric supply with uptake of this resource through their roots?

A plant which has adopted this strategy is described on page 694 of this issue of *Nature*. It is *Stylites andicola*, a member of the same pteridophyte family as the quillworts, *Isoetes*. *Stylites* grows on decomposing peat in the Andes, a carbon reserve which is definitely worth exploiting. In the opinion of Keeley, Osmond and Raven, who have examined the species both anatomically and physiologically, the thick cuticle which covers the aerial parts of the plant is essentially impermeable to CO₂ and, since there are no stomatal pores, CO₂ must enter the plant entirely through the roots. This fact was confirmed by exposing the root system to labelled ¹⁴CO₂ and showing that the rate of CO₂ incorporation was up to 50 times as fast as during the exposure of green tissues.

Ecologically, it is understandable that this type of adaptation should be found in a habitat where there is an accumulation of organic material in the soil and also where a seasonally fluctuating water table stimulates the respiratory activity of microbes and generates a rich supply of CO₂. Other peatland plants need to be

examined for this root uptake of CO₂, though none is likely to be as dependent as *Stylites* with its lack of stomata.

Keeley *et al.* also observe that, as in other members of the Isoetaceae, there is a diurnal acid metabolism of the CAM type. In the genus *Isoetes*, this characteristic is associated with an aquatic or seasonally aquatic habitat and becomes less marked when the plant is taken out of water (Keeley, J. E. *Am. J. Bot.* **69**, 254; 1982). *Stylites*, it seems, gains advantage from CAM in a rather different way. The authors imply that diurnal variation in acidity was observed in this species only in the aerial chlorophyllous tissues and it does not appear to be associated with dark short-term carbon assimilation in the

roots. A lot more needs to be known about carbon transport in the plant before this can be adequately understood.

The evolutionary significance of the discovery is considerable, particularly since it may offer some clues to the initial colonization of land by plants. Were the early colonists, like some modern aquatic plants, able to take up CO₂ through their roots? And was this capacity lost with the development of stomata? How widespread was CAM among primitive land plants, and what was its adaptive significance? Perhaps it was a hangover from aquatic environments and was later deactivated in what Van Valen (*Evol. Theory* **4**, 143; 1979) has called the 'switchback evolution' to which CAM/C4 strategies have been subjected in the course of geological time. A more detailed study of *Stylites* should throw light on at least some of these questions. □

Peter D. Moore is a Senior Lecturer in the Department of Plant Sciences, King's College, 68 Half Moon Lane, London SE24 9JF.

Astronomy

Nucleosynthesis in barium stars

from Virginia Trimble

A CLASS of stars characterized by moderate excesses of barium, yttrium, carbon and the rare-earth elements might not seem to present any dramatic feature likely to excite astronomers, but the fact is that the barium stars provide a rare opportunity to see that nucleosynthesis really is taking place in stars in our Galaxy. The anomalous elements are just those that should be produced by the s-process (slow capture of neutrons by iron nuclei) in a star whose main energy source is the burning of helium to carbon. The only trouble is, the nucleosynthesis seems to be happening in the wrong sorts of stars. The barium stars are too small and faint and too low in mass to have reached the evolutionary phase (burning of hydrogen and helium in thin shells around a carbon-oxygen core) when, according to current models, the anomalous elements ought to be produced and dredged up to the surface for us to see them. We still do not really understand how the stars do it.

Restrained rejoicing nonetheless occurred when Robert McClure and his colleagues at Dominion Astrophysical Observatory showed that essentially all the traditional barium stars are members of binary systems^{1,2}. They have separations of ~2 AU and inconspicuous companions of near 0.5 solar masses that could be, and in a few cases definitely are³, white dwarfs. The obvious solution, then, was that the nucleosynthesis had occurred in an initially rather massive star, the large separation permitting this star to expand sufficiently to reach the evolutionary stage needed to make carbon and the s-process elements

and to bring them to the surface. Then, when the massive star ejected its outer layers (in a wind and/or planetary nebula, leaving a white dwarf core behind), it polluted the surface of its previously normal neighbour. A variant scheme locates the synthesis on the surface of the white dwarf, where matter accreting from the barium star's wind can trigger nova-like flashes. These blow the material, now modified in composition, back where it came from. This would presumably happen at intervals throughout the system's life. Although the basic pollution scheme has some enthusiastic defenders⁴, recent observations⁵⁻⁹ present several difficulties for it and, instead, appear to favour an arrangement whereby the companion has somehow merely triggered mixing in the star we see.

First, the variant scheme can probably be ruled out, because the ratio of radioactive ⁹³Zr to stable ⁹³Nb, as well as the absence of unstable Tc, argues for the completion of the processing at least 3 × 10⁶ yr ago⁹. Second, a small group of stars (including such bright ones as ε Peg) shows a pattern of enhancement of only a few s-process nuclides, strongly suggesting that a very large fraction of the envelope has been exposed to only a few neutrons per seed nucleus, thereby juggling abundances within the heavy elements rather than making fresh ones out of iron⁷. It is probably not practical to transfer most of the envelope from one star to another. This is, then, an objection to the main binary model, as are the following points.

Lithium abundances present a third

difficulty. The element is fragile and easily destroyed in all but the coolest outer layers of an evolving star. Yet most barium stars have the same lithium abundance as normal K giants of the same mass (1–2 $M_{\odot}$), temperature (4,000–5,000 K), luminosity ($M_v = +2$ to -2), iron abundance (solar to 1/3 solar) and evolutionary phase (core helium burning)⁸. Thus, the supposed material to have been transferred must either have been pure carbon and s-process material (in which case so little is required that no other abundances are disturbed) or have carried with it a Li/H ratio appropriate to the mass and condition of the recipient, not the donor (supposed to be more massive and more evolved in the binary transfer model). The lithium data really only make sense if most of the present giant envelope has always been on the surface of the star, with modest amounts of pollutants convected into it. The idea that the star has been completely mixed, as already suggested¹⁰ for the still fainter CH subgiants (also carbon- and s-process-enriched), is even more at odds with the data.

A fourth difficulty arises in a detailed analysis of the ratios of abundances of 15 s-produced elements in seven assorted barium stars⁶. Different phases of stellar evolution provide different neutron sources: $^{13}\text{C}(\alpha, n)^{16}\text{O}$ during core helium burning and $^{22}\text{Ne}(\alpha, n)^{25}\text{Mg}$ during the phase with two thin shells. Both make s-process products, but with different abundance patterns, and only during the latter phase do models predict that the products should get mixed to the stellar surface. Thus, we would expect the ^{22}Ne pattern for barium stars formed by binary pollution. Instead, we seem to see the ^{13}C pattern^{6,11}.

The fifth problem is the lack of evidence for a white dwarf companion in many of the systems and for stars that have been polluted in the advertized way at any evolutionary stage except $\sim$ KOIII core helium burning^{5,6,10}. For the few barium stars where there is evidence of a white dwarf companion, it has come from excess UV flux detected by the International Ultraviolet Explorer satellite (IUE)³⁻⁵. Perhaps, then, other systems have white dwarfs that are too cool to be detected by IUE. However, white dwarfs take time to cool, and giants evolve rather rapidly. Dominy and Lambert therefore conclude that a white dwarf whose immediate predecessor has polluted the remaining star will not yet have had time to cool below IUE detection limits⁵. Böhm-Vitense *et al.* do not agree⁴ and propose instead that increasing the mass of a star by transferring material to it both speeds up and slows down its evolution. The Space Telescope should be able to ferret out any sort of white dwarf ever thought of in this context and so resolve the issue.

What about polluted stars in other evolutionary phases? A massive star will inevitably complete double shell burning

and shedding when it is good and ready, regardless of the state of its companion. Objective prism searches¹⁰ have uncovered a few old metal-poor, high-velocity stars near GO, $M_v = +4$ that are rich in carbon and s-process products, (but have failed to turn up anything at all at adjacent colours or luminosities). These are the CH subgiants mentioned above. Their discoverers interpret them as products of complete mixing at helium core flash in stars of solar mass 1. The mixing shakes them up and puts them back on the main sequence near the turn-off point for old stars like those in M67. Thus, the CH subgiants seem to be largely irrelevant to our problem, except that they ought, in due course, to evolve into something like barium giant stars. Moreover, they are not known to be binaries — an IUE survey¹² of 17 CH subgiants rules out white dwarf companions hotter than 14,000 K — and they are too rare and restricted in occurrence to give rise to most of the barium star population.

The supergiant carbon stars of types R, N and S are also not relevant. They are sufficiently bright and massive for the conventional wisdom to apply, and their s-process excesses are presumably a result of reactions in their own interiors, followed by dredge-up to the surface. Similar remarks apply to analogous older stars, the CH giants, although many of them seem to be members of binary systems¹³.

As for main-sequence stars that might have been polluted by binary companions, none is known. Yet, if 1 giant in 100 is affected (the ratio of barium stars to normal K giants), so should be 1 main-sequence star in 100 (it will capture less of the ejecta, but it also has a much smaller convective envelope to dilute what is captured). Perhaps we just have not looked carefully enough for them^{4,6}, although an awful lot of people have looked at an awful lot of main-sequence stars over the years!

Given all the difficulties with the notion of transfer, the possibility that the binary companion merely triggers unexpected mixing in the star is worth looking at. But, immediately, we hit the difficulty that the companions of the barium stars are all of low mass and inconspicuous. Yet a large fraction of binaries contain two stars of comparable mass. Thus, either a little bit of triggering works better than a lot — which seems silly — or it only works after the stars have had a chance to interact tidally through several evolutionary phases, and no effect is seen until the secondary is ascending the giant branch — which is almost as hard to believe, though perhaps not impossible⁵. There is also a difficulty of time scale 4. The s-process waits for β decays and so goes slowly. Thus, some 10^4 yr of processing are needed before the mixing event (which is assumed to be the violent phase of helium core flash). Yet an adequate neutron source will not be available until helium burning begins; and existing models indicate that violence sets

in only a few years after helium ignition¹⁴.

Further confusion is introduced by the (apparently single) K5 V star ϵ Indi. It is the closest star of its kind to us and one of the 15 closest stars altogether, with a parallax of $0.29''$. It ought, therefore, to be of the common or garden variety. Yet a detailed spectral analysis¹⁵ shows it to have excess carbon and s-process elements in a pattern very like that of the barium stars and CH subgiants (in combination with high space velocity, absolute magnitude and low iron abundance, all suggestive of an old subdwarf). Kollatzschny advised caution in accepting these results at face value. Ionized and neutral species of the same element typically gave discordant results, and the apparent anomalies in ϵ Indi may thus reflect only our incomplete understanding of the structure of very cool atmospheres. Where only one ionization state is seen, as in the group of stars studied by Holweger and Kovács⁷, improper interpretation of atmospheric structure could be responsible for all the anomalies reported¹⁵. And even for the frequently-studied bright K giant Arcturus, there are still large error bars on values for its temperature, surface gravity and abundances¹⁶.

Misinterpretation of observational data would be the easiest way to resolve the problem of apparent nucleosynthesis in barium stars, though the soul quakes at the number of wasted astronomer-hours. I do not think that this is the whole answer, but neither mass transfer from a companion nor anomalous mixing, triggered by the companion, is entirely satisfactory either. Thus, although the nuclear physics of the s-process is in good shape, we still do not seem to know its astrophysical context very well. □

1. McClure, R.D. *et al. Astrophys. J. Lett.* **238**, L35 (1980).
2. McClure, R.D. *Astrophys. J.* **268**, 264 (1983).
3. Böhm-Vitense, E. *Astrophys. J. Lett.* **239**, L79 (1980).
4. Böhm-Vitense, E., Nemec, J. & Proffitt, C. *Astrophys. J.* **278**, 726 (1984).
5. Dominy, J.F. & Lambert, D.L. *Astrophys. J.* **270**, 180 (1983).
6. Smith, V.V. *Astr. Astrophys.* **132**, 326 (1984).
7. Holweger, H. & Kovács, N. *Astr. Astrophys.* **132**, L5 (1984).
8. Pinsonneault, M.H., Sneden, C. & Smith, V.V. *Publ. astr. Soc. Pac.* **96**, 239 (1984).
9. Smith, V.V. & Lambert, D.L. *Publ. astr. Soc. Pac.* **96**, 226 (1984).
10. Luck, R.E. & Bond, H.E. *Astrophys. J.* **259**, 792 (1982).
11. Tomkin, J. & Lambert, D.L. *Astrophys. J.* **227**, 209 (1970).
12. Luck, R.E. Private communication.
13. McClure, R. *Astrophys. J. Lett.* **280**, L31 (1984).
14. Cole, P.W. & Deupree, R.G. *Astrophys. J.* **247**, 607 (1981).
15. Kollatzschny, W. *Astr. Astrophys.* **86**, 308 (1980).
16. Trimble, V. & Bell, R.A. *Q.J.R. astr. Soc.* **22**, 361 (1981).

Virginia Trimble is Professor of Physics at the University of California, Irvine, California 92717, and Visiting Professor of Astronomy at the University of Maryland, MD 20742.

Erratum

In the article 'Antibodies to the rescue' by David Hanke (*Nature* **310**, 272; 1984), in the third paragraph, auxins are described as 'plant hormones'. This description was introduced during the editing of the article; the author would prefer auxins to be referred to as 'plant growth substances'.

The meaning of homology

SIR — The recent correspondence about the word homology has been concerned more with its use (or misuse) than with its meaning. It has been generally assumed that homologous DNA sequences and protein and chromosome structures reflect either a common ancestry or convergent evolution. The lack of demonstrable similarity is held to imply either distinct ancestry or adaptive radiation¹. Yet some proteins which lack similarities may have had them and lost them through drift. More positively, one may say that the mere absence of homologies suggests nothing. Moreover, as the concurrent phylogenetic controls are not available, some accepted explanations, accounting for the existence of similarities, may ultimately require revision.

Alternatively, we could interpret the variety of homologous relations by considering the genes as determining both the evolutionary *propensities* (that is, Popper's measures of probabilities to evolve along preferred pathways under certain selection pressures²) and the critical structural features of the proteins encoded. Popper ascribes a concrete physical existence to propensities, which significantly limit the randomness of those homologies and which could not be accounted for by the above mechanisms.

Given the propensities to evolve similarly under particular evolutionary constraints, we must expect that structurally different proteins (having distinct origin) may yet display homologous amino acid sequences, whereas functionally related proteins (either of common origin or by convergence) may happen to lack them. Likewise, the propensity interpretation predicts molecules with unrelated function (and distinct origin) bearing conformational homologies, while structurally dissimilar proteins may share similar functions and have a common origin³. These predictions are supported by the factual evidence³⁻⁵, qualify a gene as a hereditary carrier of evolutionary propensity and give an unexpected twist to the old question of structure-to-function relationship. The propensity viewpoint could be tested against blind chance once the rates of evolution are known.

Moreover, the propensity interpretation of probability in evolutionary relations challenges convergence as the putative explanation of certain similarities³. This is illustrated by analogy with the karyotype evolution in chronic myeloblastic leukaemia. The terminal clinical phase of the disease is associated with the development of increasingly similar karyotypic changes of the malignant clone in more than 90 per cent of patients⁶. As the similar chromosome aberrations accumulate, an impression of "convergent" karyotype evolution is simulated. The individual patients are unrelated yet a

fictitious idea of convergence arises.

But this artefactual convergence towards a similar chromosome pattern is merely a manifestation of common but previously latent karyotype propensities to evolve in this particular blood disorder⁶. Extensive similarity of karyotype patterns does not imply genealogical (homologous) relations among the patients.

Since our extrapolations of homologies from the observed similarities often neglect the principle of maximum parsimony, we succumb unawares to a subtle prejudice that close resemblance inevitably derives from evolutionary relatedness (or chance). Unless the propensity interpretation is appreciated, such inference will continue to rationalize our sense of *déjà vu*, camouflaging it with the often vague metaphor of "convergence".

G.M. JANKOVIC

Faculty of Medicine,
University of Belgrade,
ul. Koce Todorovića bb,
11.000 Belgrade,
Yugoslavia

1. Doolittle, R.F. *Science* **214**, 149 (1981).
2. Popper, K.R. *Objective Knowledge: An Evolutionary Approach* 5th impression (Clarendon, Oxford, 1979).
3. Grütter, M.G. *et al. Nature* **303**, 828 (1983).
4. Goubin, G. *et al. Nature* **302**, 114 (1983).
5. Matthews, B.W. *et al. Nature* **290**, 334 (1981).
6. Stefanović, S. *Hematologija* 4th edn (Medicinska Knjiga Beograd, Zagreb, 1981).

Wigner's phase-space density function

SIR — I and my colleagues were surprised by the article of J. Maddox (*Nature* **308**, 601; 1984) on Wigner's phase-space density function in quantum mechanics (which, as well known, may take negative values) and by the importance attached there to a recent paper by K. Wodkiewicz. This author has in fact merely discovered in 1984, and in a particular case, a well known result: if one considers both a (one-particle) state and a measurement apparatus, represented in general by a density matrix ξ and an efficiency matrix¹ F respectively, then the phase-space integral of the product of the two Wigner functions $W_\xi(p, q)$ and $W_F(p, q)$ is always positive, being in fact just equal to the detection probability $P = \text{Tr } F\xi$ ($= |\langle \psi | \phi \rangle|^2$ if ξ and F correspond to pure states ϕ, ψ , the only case considered by Wodkiewicz). His "operational probability distribution" is just the convolution product $P(p, q) = W_\xi * W_F(p, q) = \text{Tr } F_{p, q} \xi$ obtained for a class of apparatus $F_{p, q}$ translated in phase-space in the standard way [$W_{F_{p, q}}(p', q') = W_F(p - p', q - q')$]. The above results can be found, for example, in ref. 2, or in the form just described (involving explicitly the measurement apparatus) in ref. 1. Many recent papers have rediscovered this and other simple properties of the Wigner function, not found in his 1932 work, but already obtained long ago,

for example the following result, given in that form in ref. 1 and derived there from the previous one: the Wigner function always has positive mean values in any "gaussian" box in phase-space of dimension $\geq \hbar/2$, or in other words is positive after a "gaussian $\geq \hbar/2$ smoothing", a result discovered in several papers since 1975 (including one by Connell and Wigner himself in 1981).

D IAGOLNITZER

Centre d'Études Nucléaires de Saclay,
F-91191,
Gif-sur-Yvette,
Cedex, France

1. Yvon, J. C. r. *hebd. Séanc. Acad. Sci. Paris* **11**, 311; 347 (1946).
2. Iagolnitzer, D. thesis, Paris (1967); *J. math. Phys.* **10**, 1241 (1969).

Electrical neutrality of atoms

SIR — In his excellent survey¹ on the experimental and theoretical aspects concerning the electrical neutrality of atoms, Close mentions that an unpublished result by King, based on the application of the Picard-Kessler method to H₂ and He gases, is $Q < 6 \times 10^{-21}$, where Q stands for the fractional charge difference between an electron and a proton. It should be added that a later result, also unpublished and obtained by King, applying the same technique to SF₆, puts Q at $(0 \pm 3.0) \times 10^{-23}$. Subsequently King and Dylla have reported² their determination, with an acoustic method, of the upper bound for Q as 1×10^{-21} , which apparently remains the best experimental figure we have at present.

TONG B. TANG

Institute for Molecular Science,
Myodaiji, Okazaki 444, Japan

1. Close, F. *Nature* **309**, 513 (1984).
2. Dylla H.F. & King, J.G. *Phys. Rev. A* **7**, 1224 (1973).

The rotation of birds' eggs

SIR — With reference to the interesting note by Jon Darius on the rotation of eggs (*Nature* **308**, 691; 1984), may I suggest that the blunt end of an avian egg might well have a function parallel to that of the mammalian fetal skull in stimulating the uterine muscles into contraction? Could the engaging of the blunt end be more effective in initiating the avian equivalent of "labour" than the tapered end, thus expediting the process of egg laying?

An alternative explanation of which I am less sure could relate to the development of the air sac inside the blunt end of the egg. At what stage in egg development does this air collect? Is it before laying, or is it while the egg is passing towards the cloaca? If the latter is the case, there would be an obvious advantage in presenting the egg blunt end first.

C.A. MICHIE

52 Combe Park Road,
Bath BA1 3NG, UK

Alignment of base sequences

SIR — Fitch (*Nature* 309, 410, 1984) uses gaps in aligning two base sequences "to increase the number of base matches." He possibly "misleads the careless reader" by not mentioning that there is a 25 per cent chance of two bases being identical in two unrelated sequences, each containing equal quantities of A, C, G and T. The two sequences in his first diagram have more matches (7 in 27 bases aligned) when aligned one over the other *without* a gap:

CCTTCAGAATACAGAATAGGACATAGAGA

ATCCACCCAGCCCGTGGACCTGTAT---

than with the gaps used (3 in 21 bases aligned). In Fitch's second diagram, he contrives 7 matches in 27 bases aligned by inserting a gap, which is the same amount of matching as without a gap, and the same as by random coincidence.

Any two random sequences of A, C, G and T may be made to give the appearance of matching if enough gaps are arbitrarily inserted. Each such gap quantitatively diminishes the probability of evolutionary homology¹. Obviously, there is no point in inserting gaps that do not increase the number of matches, or in using algorithms to no purpose.

THOMAS H. JUKES
Space Sciences Laboratory,
University of California,
Berkeley, California 94720, USA

1. Jukes, T.H. & Cantor, C.R. *Mammalian Protein Metabolism* Vol. 3, 21-132 (Academic, New York, 1969).

Opiates and sexual function

SIR — We were interested to note that Weisenfeld-Hallin and Sodersten¹ were able to confirm (in rats) previous findings in man that spinal opiates interfere with male sexual function. In 1980, Torda *et al.*² reported failure to ejaculate in three out of four sexually experienced Australian anaesthetists who had been pretreated with lumbar extradural morphine. We feel stimulated to enlarge on this; although we do not have precise measurements of the number of intromissions or the duration of intercourse in the latter study, the incidence of ejaculatory failure was in fact 100 per cent, as the fourth volunteer was not introduced to a sexually receptive female within the expected duration of effect of the extradural injection.

It may be true² that spinal opiate research in laboratory animals has concentrated exclusively on pain perception, but this is not the case in man. There is now abundant evidence that spinal opiates interact with a far wider variety of neural pathways than originally thought. The endocrine responses to nociceptive stimuli^{3,4}, bladder function², psychogalvanic responses⁵ and respiratory function^{4,6} are all modified by spinal opiates.

Spinal opiates appear to affect sexual

performance primarily by modifying sympathetic responses to sexual stimulation. Whereas the efferent pathways for penile erection in man are thought to be parasympathetic, ejaculation and termination of the erection are believed to be under the control of sympathetic efferents. One might predict, therefore, that sympathetic blockade by a spinal opiate would lead to sustained erection with ejaculation, as has been observed in man.

Naloxone reversal of sexual dysfunction has not been investigated in man, but its occurrence in rats implies that the effect is specific to opiates. This may not be so, as we have observed delayed, but ultimately successful, ejaculation after the use of caudally administered extradural steroids. That the sexual effects of spinal opiates may be nonspecific is not altogether surprising; concentrations of morphine in spinal fluid after intrathecal injection may be over 100,000 nmol per litre⁷, and a significant fraction of extradural morphine may cross the dura⁸ to give very large spinal fluid drug concentrations. These are far in excess of concentrations required for saturation of spinal opiate receptors.

Whatever the mechanism, there is at least one important clinical implication for these findings. Spinal administration of opiates or other drugs may constitute an important new therapeutic tool in the treatment of premature ejaculation.

A. PYBUS
T. TORDA

Department of Anaesthesia,
Prince Henry Hospital,
Sydney, Australia

H.J. MCQUAY

Nuffield Department of Anaesthetics,
Abingdon Hospital,
Abingdon, Oxford, UK

R.A. MOORE

Nuffield Departments of Clinical
Biochemistry and Anaesthetics,
Radcliffe Infirmary, Oxford, UK

1. Weisenfeld-Hallin, Z. & Sodersten, P. *Nature* 309, 257-258 (1984).
2. Torda, T. *et al. Br. J. Anaesth.* 52, 939-942 (1982).
3. Cowen, M. *et al. Anesth. Analg.* 61, 15-18 (1982).
4. Moore, R.A. *et al. Br. J. Anaesth.* (in the press).
5. Leslie, J. *et al. Lancet* i, 151 (1979).
6. Rawal, N. & Wattwil, N. *Anesth. Analg.* 63, 8-14 (1984).
7. Moore, R.A. *et al. Clin. Pharm. Ther.* 35, 40-46 (1984).
8. Moore, R.A. *et al. Br. J. Anaesth.* 54, 1117-1128 (1982).

Adiabatic boiling and ocean metal transport

SIR — R.N. Anderson¹, commenting on Delaney and Cosens², presents a misconception about adiabatic boiling, as well as an error about precious metal transport and deposition. Both articles assume that adiabatic ascent of heated seawater beneath thermally active ocean ridges³ can result in a greatly concentrated residual liquid due to steam loss if the liquid intersects and follows the subcritical temperature-pressure two-phase curve. This is their explanation for fluid inclusions with salinities two to five times that of

seawater in seafloor basalts².

The dissolved salts in seawater raise its critical point above that of pure water to about 405°C⁴. Using thermodynamic data for a seawater-equivalent NaCl solution⁵, and enthalpy calculations by K.S. Pitzer (personal communication), one may calculate the maximum steam fraction that will result due to adiabatic boiling from 400°C and, say, 500 bars³ to 200°C (the maximum temperature range over which boiling might occur before magmatically heated seawater is expelled from an ocean ridge). About 53 per cent of the original mass of the liquid would be converted to steam, thus concentrating the salinity of the original to the residual liquid by a factor of about 1.9.

I do not argue that boiling cannot occur beneath the ocean floor. It may, particularly beneath some of the shallower ridges, or where the gas contents are high. But the concentration factors cited² cannot be attributed solely to boiling, and the fluid inclusion data are not conclusive evidence of boiling. A possible mechanism for the creation of the brines in fluid inclusions may be supercritical phase separation of seawater^{4,6}.

Anderson also claims that Delaney and Cosens "predict the steam" generated by such boiling may carry gold and silver to shallower parts of a hydrothermal system, leaving lead, zinc and copper below the zone of boiling. However, Delaney and Cosens² do not suggest (nor does anyone else) that precious metals are transported by subcritical steam. Rather, it is simply observed in some ore deposits that precious metals are often stratified above base metals, with the point of separation being the zone of first boiling. Although vapour may be present above this zone, it is the residual liquid that continues to be the transporting medium of the non-volatile species. This is the situation in the geothermal systems of New Zealand, where ore-grade gold and silver precipitates are being formed⁷.

Studies of active hydrothermal systems and their hot springs are very important to understanding how fossil hydrothermal systems formed ore deposits; they may assist in finding new deposits and even new environments of mineralization. But Anderson's speculations, particularly on finding a deep sea "steam vent where we might observe precious metal deposits forming *in situ* for the first time ever", are not justified.

JEFFREY W. HEDENQUIST

DSIR Geothermal Research Centre,
Wairakei, Private Bag,
Taupo, New Zealand

1. Anderson, R.N. *Nature* 303, 283 (1983).
2. Delaney, J.R. & Cosens, B.A. *Mar. Technol. Soc. J.* 62, 62-66 (1982).
3. Bischoff, J.L. *Science* 207, 1465-1469 (1980).
4. Bischoff, J.L. & Rosenbauer, R.J. *Earth planet. Sci. Lett.* (in the press).
5. Pitzer, K.S. & Li, Y.-G. *Proc. natn. Acad. Sci. U.S.A.* 80, 7689-7693 (1983).
6. Delaney, J.R. *et al. EOS* 63, 472 (1982).
7. Weissberg, B.G. *Econ Geol* 64, 95-108 (1969).

Can a neutrino-dominated Universe be rejected?

Piet Hut* & Simon D. M. White*

Department of Astronomy, University of California, Berkeley, California 94720, USA

The conservative hypothesis that the Universe is dominated by nonbaryonic matter made up of particles that have already been observed in the laboratory is not tenable within a standard big-bang cosmology if all neutrino species are stable. It can be salvaged only if one neutrino species has a mass in the range 0.1–250 MeV and decays on a short mass-dependent time scale, while a second species has a mass in the range 10–50 eV and is stable.

THE masses of large astronomical systems, such as groups and clusters of galaxies, can be determined from the requirement that gravitational forces be sufficient to induce the observed random motions of their constituent galaxies. This procedure is directly analogous to determining the mass of the Sun from its distance from the Earth and from the orbital velocity of the Earth. Masses found in this way exceed by more than an order of magnitude the total mass of the stars in the visible parts of the galaxies. A similar discrepancy shows up when the motion of stars and gas in the outer parts of the disks of spiral galaxies is used to analyse the distribution of mass within them. Most of the mass seems to be located far from the centre, well outside the regions which contain most of the light. This difficulty has been dubbed the 'missing mass problem' although, in fact, it is the light which seems to be missing—the mass is certainly there (see ref. 1). It seems that at least 90% of the mass of the Universe is in some form which is dark and has so far remained undetected except by its gravitational effects.

Many candidates have been suggested for the dominant dark component, spanning an embarrassingly large mass range of 78 orders of magnitude. Some possibilities are: black holes (10^{15} – 10^{40} g); baryonic material in the form of small stars, planets or rocks ($<10^{32}$ g); elementary particles such as monopoles ($\sim 10^{-7}$ g), photinos ($\sim 10^{-30}$ g), neutrinos ($\sim 10^{-31}$ g) and axions ($\sim 10^{-38}$ g). Of these candidates, only stars, planets, rocks and neutrinos have been directly observed and a nonzero rest mass has not yet been confirmed for any neutrino. The suitability of the other candidates is doubly uncertain: they must be shown to exist in nature, as well as to have a mass which matches their present number density in such a way as to produce the observed cosmic matter density.

The axion, for example, is the particle corresponding to a hypothetical quantum field originally suggested as an explanation of why processes that break charge–parity symmetry occur much more rarely in nature than seems natural in the standard model for the strong interactions. Soon after this possibility was proposed, it was realized that the axion would need to have a mass $<10^{-7}$ electron masses, if it was not to adversely affect the energy loss processes from the centres of stars but $>10^{-11}$ electron masses, if it were not to be so copiously produced in the big-bang that the total density of axions would greatly exceed the present density of the Universe^{36–39}. Fortunately, as well as providing the present limits on axion masses, astrophysics also provides sufficiently abundant sources of axions that there is some chance of detecting them directly on Earth despite their

incredibly weak interactions. Nevertheless, there is no direct experimental evidence for their existence.

In view of the difficulty of distinguishing between the plethora of exotic candidates proposed for the dark matter, it seems wise to concentrate first on objects which are known to exist. A powerful constraint comes from the theory of nucleosynthesis in the standard big-bang cosmology. Although the amount of mass inferred from observations of galaxy clustering and of the outer parts of spiral galaxies is just marginally consistent with the hidden matter being made of baryons, the observed element abundances are not consistent with a present baryon density sufficient to close the Universe and to ensure that its expansion will one day be arrested. However, inflationary models for the early Universe² and general philosophical considerations suggest that the present total density should be extremely close to that required for closure, otherwise the fact that it does not differ from this value by many orders of magnitude would need to be ascribed to some improbable cosmic coincidence. Retention of a flat Universe and of standard cosmology requires that the dark matter cannot be baryonic; the only such candidates known to exist are the three types of neutrino (ν_e , ν_μ and ν_τ). In this article we discuss current constraints on a neutrino-dominated Universe.

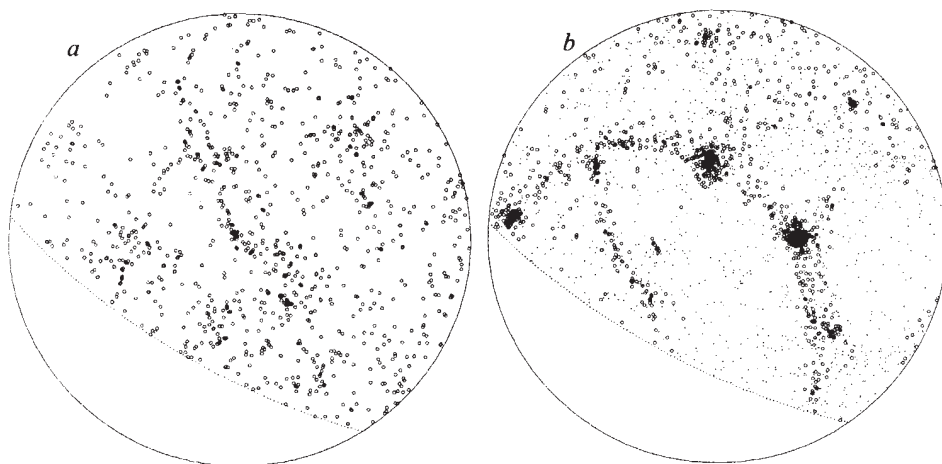
Stable massive neutrinos

While the simplest theories for the unification of the strong and electroweak interactions (that based on the symmetry group SU(5)) requires neutrinos to be massless, many other grand-unification schemes allow them to acquire a nonzero rest mass. Unconfirmed claims that a nonzero neutrino rest mass had been detected^{3,4} stimulated interest both in such theories and in neutrino-dominated cosmologies. Provided their lifetimes are much greater than the age of the Universe, the role of massive neutrinos in cosmology is relatively simple and depends only on their masses. (We do not discuss the possibility of a significant nonzero chemical potential for the neutrino distribution; besides being rather unnatural it does not weaken the constraints we describe.) We consider this simple case first; the present Universe is then dominated by the most massive stable neutrino species and we assume that no other species has decayed since neutrino decoupling ($kT \sim 2$ MeV).

The present number density of each species of neutrino is $\frac{3}{4} \times \frac{4}{11}$ times the number density of the photons that we observe in the microwave background radiation; the first factor comes from the difference in statistics and the second from the additional production of photons by e^+e^- annihilation⁵. The mass of the dominant particle is thus very simply related to the present mean cosmic density. This mass also determines the epoch and thus the horizon scale of the Universe when the particles became nonrelativistic.

* Present addresses: Institute for Advanced Study, Princeton, New Jersey 08540, USA (P.H.); Stewart Observatory, University of Arizona, Tucson, Arizona, USA (S.D.M.W.).

Fig. 1 *a*, An equal area projection of the northern sky. The outer circle is at galactic latitude $+40^\circ$ and the dashed line corresponds to 0° declination. Symbols show those galaxies with apparent magnitude $m_z \leq 14.5$ and recession velocities $v_r < 10,000 \text{ km s}^{-1}$ which would still have apparent magnitude $m_z \leq 14.5$ if their distance were scaled in proportion to $v_r/4,000$. This catalogue³⁴ is thus volume limited to $4,000 \text{ km s}^{-1}$ and magnitude limited beyond that. *b*, An N -body model of a neutrino-dominated universe⁹ in which structure first formed at a redshift of 2.5. The model was scaled to $\Omega = 1$, $H_0 = 50 \text{ km s}^{-1}$, $m_\nu = 24 \text{ eV}$ and particles were plotted using the same selection in depth as applies to the real data. Points in collapsed regions where galaxies could, in principle, be formed are shown as open circles, whereas points in uncollapsed regions are shown as dots and cannot represent galaxies. The number of open circles is the same in *a* and *b*. Note the very large empty regions expected in the neutrino-dominated Universe.



This horizon scale is, in turn, related to the minimum scale of spatial variations in the density of neutrinos at later epochs. When a neutrino-density fluctuation is encompassed by the particle horizon, streaming of neutrinos away from the peak will reduce its amplitude drastically if the neutrinos are still relativistic. Large-scale fluctuations enter the horizon after the neutrinos have slowed down and are not significantly affected by this process. Plane-wave density fluctuations of current wavelength λ_d are reduced in amplitude by an order of magnitude, where $\lambda_d = 41/(m_{30}\vartheta) \text{ Mpc} = 13/(\Omega h^2 \vartheta^{-2}) \text{ Mpc}$ (refs 6–8). (m_{30} here is the mass of the dominant neutrino species in units of 30 eV, ϑ is the microwave background temperature in units of 2.7 K, h is the Hubble parameter in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and Ω is the cosmic mass density in units of the value required to close the Universe.) This is a very large scale, comparable with the largest structures seen in the present Universe; smaller fluctuations in the neutrino distribution were strongly damped. Detailed calculations^{6–8} of the linear transfer function for this process specify the statistical properties of the neutrino distribution at redshifts $z \geq 10$ for any assumed density fluctuation spectrum at early times. The shape of the initial fluctuation spectrum is unambiguously determined in inflationary models for the generation of structure but its overall amplitude is model dependent and uncertain; this amplitude determines the epoch when nonlinear structures ('pancakes') first form.

Numerical simulations of the nonlinear growth of structure from initial fluctuation spectra of this kind show that $\lambda_d < 8/h \text{ Mpc}$ is required for consistency with the observed scale of galaxy clustering and with the requirement that structures begin forming by a redshift > 4 (ref. 9). This limit cannot be reconciled with the value expected in a neutrino-dominated Universe for any acceptable cosmological parameters; both the Soviet measurement of the electron neutrino mass and the requirement that the age of the Universe be greater than that of globular star clusters ($> 12 \text{ Gyr}$) lead to disagreement by a factor of at least 3 in length scale (for $\Omega < 1$). The conflict is exacerbated if several neutrino species have the same mass. (The possibility $m_\nu > 3 \text{ GeV}$ (refs 10–13) can lead to a reasonable value for the coherence length¹⁴, but is inconsistent with experimental limits on the masses of all three known neutrinos¹⁵.)

Figure 1 illustrates the scale discrepancy by comparing the observed distribution of galaxies in the northern sky with the distribution of matter in a simulation of a neutrino-dominated Universe scaled assuming that $m_{30} = 0.8$, $h = 0.5$, $\Omega = 1$ and that 1% of matter had undergone local collapse by a redshift of 2.5. The inconsistency between the two pictures arises not because λ_d in a neutrino-dominated Universe is much larger than any structures seen in the real world but rather because if structure

first forms at $z > 2.5$ the present density contrast on scales $\sim \lambda_d$ is predicted to be much higher than is observed. Galaxy formation in a neutrino-dominated Universe is expected to occur only in regions which have undergone local collapse. Figure 1 shows that this requirement seems likely to make the galaxy distribution even more inhomogeneous on large scales than the neutrino distribution. It seems that a neutrino-dominated Universe can be ruled out if all three neutrino species have been stable since decoupling.

Unstable massive neutrinos

We must now enquire whether decay of a massive neutrino can allow us to evade the above conclusion without the introduction of exotic physics or cosmology. Our approach is to find restrictions which are independent of any assumed model for neutrino decay. For example, we include restrictions on heavy neutrinos formed in the 'heat bath' of a supernova explosion¹⁶, but we exclude restrictions on heavy neutrinos which follow from the presence of the energetic tail of the ν_β flux from the Sun, assuming a particular type of coupling involving a small mixing angle¹⁷. Although the latter restriction is based on a very reasonable assumption about neutrino mixing, that assumption is, nevertheless, model dependent. We also exclude the possibility that ν decay may occur on a much shorter time scale than the minimum predicted by standard weak interaction theory. While such decay might have cosmologically interesting consequences^{18,19}, it requires the mediation of exotic quantum fields (such as, familons or majorons) for which there is no experimental support. In addition, such schemes are not consistent with our conservative assumptions because they predict that almost half the present energy density of the Universe should be contributed by these exotic particles.

If the decay is radiative it can increase the density of photons relative to that of the less massive, stable neutrinos which currently dominate the Universe. The latter may then have a higher mass and hence a smaller initial coherence length without conflicting with limits on the mean cosmic density. If neutrino decay is nonradiative, it may also be possible to escape our previous conclusion: the unstable particle or its decay products can enhance the energy density of the Universe at the time the stable neutrino becomes nonrelativistic. A higher energy density requires a higher expansion rate to keep the Universe flat at this early time, and therefore a smaller age and a smaller horizon size. This effect lowers the coherence length imposed on the distribution of light particles and may eliminate any conflict with the observed scale of galaxy clustering. Davis *et al.*²⁰

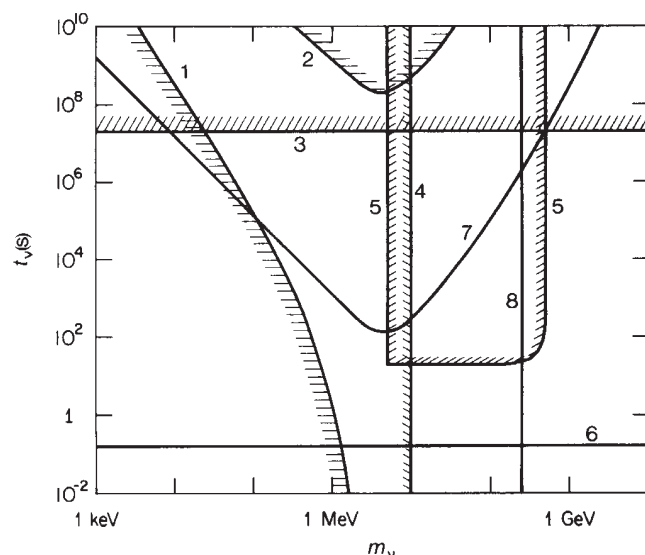


Fig. 2 Constraints on the mass m_ν and lifetime t_ν of radiatively-decaying heavy neutrinos from: 1, $\nu_e - e$ scattering experiments²²⁻²⁵; 2, the present density of the Universe²²⁻²⁵; 3, thermalization of decay photons in the background radiation³⁵; 4, supernova observations¹⁶; 5, deuterium destruction²¹ (caveat, see text). The first scenario (see text) requires a lifetime above line 6 (refs 22-25), where neutrinos decouple, and above line 7, to increase the present photon number density by a factor of 3 (as is required to reduce the present coherence length by the same factor). Line 8 shows the upper limit on the mass of ν_τ from laboratory experiments¹⁵.

identified the first of these possibilities and explored a complex special case of the second. For simplicity, we assume natural neutrino mass and lifetime ordering ($m_\tau > m_\mu > m_\nu$). In addition, call the unstable particle ν_τ and the particle which currently dominates the Universe ν_μ .

Radiative decay, $\nu \rightarrow \nu + \gamma$, is subject to a variety of cosmological and laboratory constraints which we show in Fig. 2; only the bottom right-hand corner is allowed. Unfortunately, neutrinos in this region do not live long enough to produce the increase in photon temperature that we require. The present neutrino temperature cannot be low enough to allow $m_\nu > 85$ eV as is needed to get a sufficiently small coherence length. Our analysis differs slightly from that of Davis *et al.*²⁰ who concluded that this scenario could not produce galaxy haloes, but that, for $m_\nu \sim 10$ MeV it could produce coherence lengths of the size that we require. We exclude this latter possibility too, since for this mass the neutrino lifetime t_ν must be less than a tenth the age of the Universe at the time of nucleosynthesis, otherwise the decay photons destroy too much deuterium²¹. At this early time (10-20 s) the neutrinos do not yet dominate the energy density of the Universe, and cannot significantly increase the photon temperature. There is one caveat here: Fig. 2 shows the excluded area determined by Lindley²¹, although he overlooked the possibility that decay photons might lose energy by producing e^+e^- pairs in collisions with background photons. This weakens his restrictions significantly in the region of Fig. 2 where the centre of mass energy of such collisions exceeds $2m_{ec^2}$. This effect clears part of the bottom right-hand side of the excluded area 5 in Fig. 2, but the area above line 7 remains excluded. (A more exact treatment should include multiple scatterings which Lindley left out; this would enlarge the excluded area.)

Nucleosynthesis constraints from the observed deuterium abundance²²⁻²⁵ are not indicated in Fig. 2 because these do not significantly limit the parameter space of interest (to produce the observed deuterium abundance implies an upper limit on $t_\nu(m_\nu)$ which is similar in shape to line 7 but displaced towards higher t_ν values by more than two orders of magnitude). The case, $m_\nu \sim 500$ MeV, might also seem to be marginally allowed,

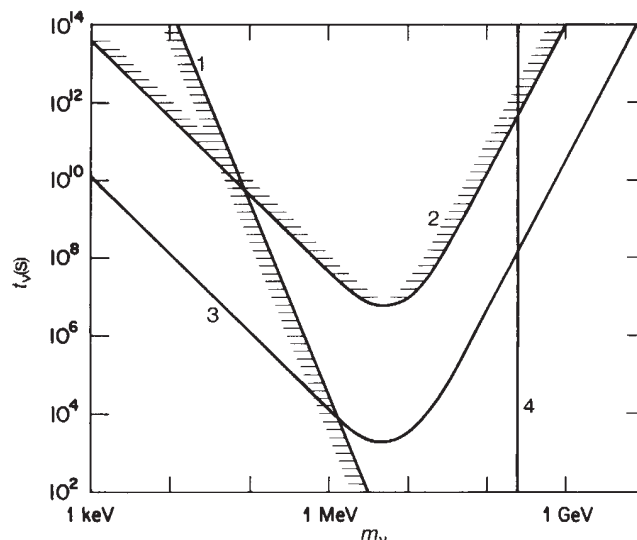


Fig. 3 Constraints on the mass and lifetime of nonradiatively-decaying neutrinos, from: 1, requiring $q > 1$; 2, requiring $1 + z_{eq} > 5$, the redshift when quasars formed—this is a very conservative limit; under plausible assumptions the large-scale isotropy of the microwave background implies a more stringent upper limit which is parallel to, and lies only slightly above, line 3. The lifetime must lie above line 3 if the horizon scale at z_μ is to be decreased by a factor of 3. Line 4 shows the experimental upper limit on the mass of ν_τ (ref. 15). Note that at the intersection of lines 3 and 4, $q = 5 \times 10^{15}$ showing a decay rate far below that obtained by scaling from muon decay.

as the number density of neutrinos is then so small that deuterium destruction becomes tolerable²¹. However, three arguments appear to rule out this possibility: (1) multiple scattering of decay photons makes deuterium destruction more efficient than in the calculations reported in Fig. 2; (2) nonequilibrium corrections to decay rates significantly raise line 7 in Fig. 2; (3) line 7 is also raised by the decay channels which open up when mesons can be produced; these channels were not considered by Dicus *et al.*²²⁻²⁵.

Nonradiative decay, $\nu_\tau \rightarrow 2\nu + \bar{\nu}$ offers another means to lower the coherence length of the present neutrino distribution. Massive ν_τ s become nonrelativistic at the redshift when $m_\nu c^2 = kT_\nu$; this redshift is $1 + z_\tau = 6.0 \times 10^9 \vartheta^{-1} (m_\tau/1 \text{ MeV})$. Soon thereafter they dominate the energy density and the Universe remains matter dominated until its age equals their exponential decay time, t_ν ; the redshift corresponding to this time is $1 + z_d = 3.7 \times 10^{11} (fM_\nu \Omega h^2 / m_\mu)^{-0.33} (t_\nu/1 \text{ s})^{-0.67}$, where f is the fraction of the present cosmic density made up of 'primordial' ν_μ s and $F < 1$ is the factor by which annihilation depletes the abundance of ν_τ s before they decouple. $F = 1$ for $m_\tau < 1$ MeV; for higher energies we use the F values given in refs 22-25 although recent work by Krauss²⁶ has shown that they may need to be increased by factors of up to 6 if the neutrinos have majorana masses. This increase is insufficient to effect our results significantly. Setting $t_\nu = 2.9 \times 10^4 q (m_\tau/1 \text{ MeV})^{-5}$ s so that q is the factor by which t_ν exceeds the value obtained by scaling from muon decay, and using $m_\mu = 96 f \Omega h^2 \vartheta^{-3}$ eV, the decay redshift can be written, $1 + z_d = 1.8 \times 10^7 F^{-0.33} \vartheta^{-1} q^{-0.67} (m_\tau/1 \text{ MeV})^3$. Weak interaction theory suggests that q substantially exceeds unity.

After the ν_τ s decay the Universe is again radiation dominated and the energy density in decay products exceeds that in photons and 'primordial' ν_τ s and ν_μ s by a factor $14 F^{1.33} q^{0.67} (m_\tau/1 \text{ MeV})^{-2}$. The age of the Universe when the primordial ν_μ s go nonrelativistic (at $1 + z_\mu = 5.8 \times 10^5 f \Omega h^2 \vartheta^{-4}$) is reduced from the value in a standard ν -dominated Universe by the square root of this factor and the mass of ν_μ s within the horizon at that time is reduced by its 3/2 power; this is the reduction in coherence scale we are looking for. The energy density of the

primordial ν_μ s does not exceed that of the ν_τ decay products until the rather recent redshift, $1+z_{\text{eq}} = 3.6 \times 10^{-3} F^{-1.33} q^{-0.67} (m_\tau/1 \text{ MeV})^2 (1+z_\mu)$. (If $m_\tau < 5 \text{ MeV}$ and some of the decay products are ν_μ s, the latter go nonrelativistic at $z \approx z_{\text{eq}}$ and thereafter make a significant contribution to the overall density. Below we shall ignore this special case, which is less favourable for galaxy formation, and set $f=1$.)

The above considerations apply for the situation where $\nu_\tau \rightarrow 2\nu + \bar{\nu}$ is the only decay channel. This will be the case for purely leptonic decay at moderate energies, but it is unlikely to hold when $m_\tau > 2m_e$, since the extra channel $\nu_\tau \rightarrow \nu + e^+ + e^-$ becomes energetically accessible. For completeness as well as generality, however, we will discuss first the case in which $\nu_\tau \rightarrow 2\nu + \bar{\nu}$ is the only channel both above and below 1 MeV. We will then discuss the perhaps more realistic case when charged decay products appear above 1 MeV. Of course, in the latter case, radiative decay will take place just below 1 MeV through virtual pair production as a one-loop higher order correction, in which $\nu \rightarrow \nu + (e^+e^- \rightarrow \gamma)$. However, for energies significantly below 1 MeV this process will be strongly suppressed, and therefore introduce only small corrections to our main conclusions.

Figure 3 gives constraints on m_τ and t_ν when the heavy neutrino decays into other neutrinos only. We impose a very conservative upper bound on t_ν by requiring that the Universe should be no longer radiation dominated at the epoch when the first nonlinear structures formed. A lower limit comes from requiring a sufficient reduction in the coherence length. The range of acceptable parameters can be narrowed further using limits on the anisotropy of the microwave background radiation. For the constant curvature fluctuation spectrum predicted by inflationary models²⁷⁻²⁹, application of the methods of Peebles³⁰ shows that the observed upper limit on correlated temperature fluctuations at 10° (ref. 30) requires $1+z_{\text{eq}} > 840$ for $\Omega=1$; a more restrictive limit applies if $\Omega < 1$. This constraint lowers the upper limit of Fig. 3 by a factor of at least 2,000 and brings it near to coincidence with the lower bound (line 3). Less stringent limits on z_{eq} are found if the initial fluctuation spectrum is assumed to contain less power on large scales than a constant curvature spectrum.

In the gravitino-dominated universes for which Bond *et al.*³¹ present detailed calculations, the ratio $(1+z_\mu)/(1+z_{\text{eq}})$ takes the value 506; we find the same value if $F^{-1.33} q^{-0.67} (m_\tau/1 \text{ MeV})^2 = 0.55$ (which is plausible for any $m_\tau > 0.75 \text{ MeV}$). For such parameters the density fluctuation spectra in neutrino-dominated universes will be identical in shape to those they give, but with characteristic mass scales increased in proportion to $z_{\text{eq}}^{-1.5}$ (a factor of 102 independent of cosmological parameters). For $\Omega=1$ these spectra are consistent with all the constraints given above but they lead to $z_{\text{eq}} \leq 339$ for $\Omega h^2 \vartheta^{-4} \leq 0.3$ and so violate the background radiation constraint just discussed.

Let us now consider what will happen if $\nu_\tau \rightarrow \nu + e^+ + e^-$ becomes available as a decay channel for $m_\tau > 1 \text{ MeV}$. In this case, Fig. 3 applies for $m_\tau < 1 \text{ MeV}$, while Fig. 2 applies with some modification for $m_\tau > 1 \text{ MeV}$. The main difference is that deuterium destruction will be less efficient than indicated by

line 5 in Fig. 2, because direct photofission now has to be replaced by two other mechanisms. The first mechanism is the spallation of deuterium by a direct decay product (e^+ or e^-), and is less efficient because these particles will lose energy rapidly through inverse Compton scattering off the ambient photons. The second mechanism is photofission of deuterium by background photons which have been Compton scattered to high energy by the decay products. To reach the spallation threshold energy in a single inverse Compton scattering event becomes increasingly difficult with time, because the necessary energy amplification factor for a typical background photon increases. The resulting requirement is $[t_\nu/1 \text{ s}] < 0.3 \times [m_\tau/1 \text{ MeV}]^4$. This line happens to coincide closely with the right-hand side of line 3 in Fig. 3.

In addition to the $\nu e^+ e^-$ decay channel, there will be higher-order contributions to the $\nu\gamma$ decay channel, where only a virtual e^+e^- pair is produced, which immediately decays into a photon. These energetic photons again threaten deuterium survival, although the rate of deuterium spallation is less than in the purely radiative case by the ratio of the two decay channels mentioned above. To summarize the leptonic decay case, we can take Fig. 3 and exclude the whole region $1 < m_\tau < 10 \text{ MeV}$ (see line 4 in Fig. 2). The possibly allowed regions left over are a small, perhaps nonexistent, triangle at the left of this excluded band, and a region to the right. This last region is bounded above by line 2, or a lowered version of it, below by line 3, and to the right by line 4. It is bounded on the left by line 4 from Fig. 2, or by a more stringent constraint if deuterium spallation by energetic decay photons is sufficiently effective.

Conclusions

The dark matter in the Universe is thought to be non-baryonic. If it is made of particles known to exist from laboratory experiments, and if the standard big-bang cosmology is correct, there must be two massive neutrino species. The heavier must have a mass in the range $0.1 \text{ MeV} < m < 0.25 \text{ GeV}$ and must undergo decay on a time scale in the range shown in Fig. 3. Upper limits on large-scale anisotropy in the microwave background reduce this range to a narrow strip in the neighbourhood of its lower limit if one requires a constant curvature initial fluctuation spectrum. The lighter neutrino must have a mass of order 30 eV and must have a very long lifetime ($t_\nu > 10^{17} \text{ s}$ ($\nu \rightarrow \nu\nu$) or $t_\nu \gg 10^{22} \text{ s}$ ($\nu \rightarrow \nu\gamma$) [refs 32, 33]). Cosmologies in which all neutrino species are stable give rise to coherent structures which are too large to be consistent with the observed galaxy distribution. There seems to be no natural reason why the parameters of real neutrinos should lie in the narrow ranges required to get a viable cosmology. If they do lie in these ranges a neutrino-dominated cosmology is consistent with overall scaling constraints; however, a detailed study of galaxy formation might still show it to be inconsistent with observation.

We thank Marc Davis, Sheldon Glashow, Nick Kaiser, Lawrence Krauss, Rashid Sunyaev and Frank Wilczek for helpful comments. This work was supported by D.O.E. contract DE-ATO 3-82ER40069, and by NSF grant PHY82-17352.

- Faber, S. M. & Gallagher, J. S. *Astr. Astrophys.* **17**, 135-187 (1979).
- Schramm, D. N. *Phil. Trans. R. Soc. A* **307**, 43 (1982).
- Lyubimov, V. A., Novikov, E. G., Nozik, V. Z., Tretyakov, E. F. & Kozik, V. S. *Phys. Lett.* **B94**, 266-268 (1980).
- Reines, F., Sobel, S. W. & Pasierb, E. *Phys. Rev. Lett.* **45**, 1307-1311 (1980).
- Weinberg, S. *Gravitation and Cosmology* (Wiley, New York, 1972).
- Bond, J. R., Efstathiou, G. & Silk, J. *Phys. Rev. Lett.* **45**, 1980-1984 (1980).
- Bond, J. R. & Szalay, A. S. *Astrophys. J.* **274**, 443-468 (1983).
- Peebles, P. J. E. *Astrophys. J.* **258**, 415-424 (1982).
- White, S. D. M., Frenk, C. S. & Davis, M. *Astrophys. J. Lett.* **274**, L1-L5 (1983).
- Hut, P. *Phys. Lett.* **69B**, 85-88 (1977).
- Lee, B. W. & Weinberg, S. *Phys. Rev. Lett.* **39**, 165-168 (1977).
- Sato, K. & Kobayashi, M. *Prog. theor. Phys.* **58**, 1775-1789 (1977).
- Vysotskii, M. I., Dolgov, A. D. & Zel'dovich, Ya. B. *JETP Lett.* **26**, 188-190 (1979).
- Peebles, P. J. E. *Astrophys. J. Lett.* **263**, L1-L5 (1982).
- Bacino, W. *et al. Phys. Rev. Lett.* **42**, 749-752 (1979).
- Falk, S. W. & Schramm, D. N. *Phys. Lett.* **79B**, 511-513 (1978).
- Toussaint, D. & Wilczek, F. *Nature* **289**, 777-778 (1981).
- Turner, M. S., Steigman, G. & Krauss, L. M. Preprint, Fermilab-Pub. 84/32-A (1984).
- Gelmini, G. & Schramm, D. N. Preprint (University of Chicago, 1984).
- Davis, M., Lecar, M., Pryor, C. & Witten, E. *Astrophys. J.* **250**, 423-431 (1981).
- Lindley, D. *Mon. Not. R. astr. Soc.* **188**, 15P-19P (1979).
- Dicus, D. A., Kolb, E. W. & Teplitz, V. L. *Astrophys. J.* **221**, 327-341 (1978).
- Dicus, D. A., Kolb, E. W., Teplitz, V. L. & Wagoner, R. V. *Phys. Rev.* **D17**, 1529-1538 (1978).
- Gunn, J. E., Lee, B. W., Lerche, I., Schramm, D. N. & Steigman, G. *Astrophys. J.* **223**, 1015-1031 (1978).
- Miyama, S. & Sato, K. *Prog. theor. Phys.* **60**, 1703-1716 (1978).
- Krauss, L. M. *Phys. Lett.* **128B**, 37-43 (1983).
- Hawking, S. W. *Phys. Lett.* **115B**, 295-297 (1982).
- Guth, A. & Yi, P. S. *Phys. Rev. Lett.* **49**, 1110-1113 (1982).
- Starobinski, A. A. *Phys. Lett.* **117B**, 175-178 (1982).
- Fixsen, D. J., Cheng, F. S. & Wilkinson, D. T. *Phys. Rev. Lett.* **50**, 620-622 (1983).
- Bond, J. R., Szalay, A. S. & Turner, M. S. *Phys. Rev. Lett.* **48**, 1636-1639 (1982).
- Kimble, R., Bowyer, S. & Jacobsen, P. *Phys. Rev. Lett.* **46**, 80-83 (1981).
- Stecker, F. W. & Brown, R. W. *Astrophys. J.* **257**, 1-9 (1982).
- Davis, M., Huchra, J., Latham, D. W. & Tonry, J. *Astrophys. J.* **253**, 423-445 (1981).
- Danese, L. & De Zotti, G. *Astr. Astrophys.* **107**, 39-42 (1982).
- Preskill, J., Wise, M. & Wilczek, F. *Phys. Lett.* **120B**, 127-132 (1983).
- Abbott, L. & Sikivie, P. *Phys. Lett.* **120B**, 133-136 (1983).
- Dine, M. & Fischler, W. *Phys. Lett.* **120B**, 137-141 (1983).
- Fukugita, M., Watamura, S. & Yoshimura, M. *Phys. Rev. Lett.* **48**, 1522-1525 (1982).

Coesite in clinopyroxene in the Caledonides and its implications for geodynamics

David C. Smith

Laboratoire de Mineralogie, CNRS Laboratoire-Associé No. LA286, Muséum National d'Histoire Naturelle, 61 Rue de Buffon, 75005 Paris, France

Coesite is reported for the first time from the Caledonide orogen. It occurs as inclusions in clinopyroxene in the dolomite-eclogite at Grytting, Norway, and provides valuable new evidence to support the hypothesis of extremely high pressure (≥ 30 kbar) in certain Norwegian eclogites. This finding has implications for the geodynamics and kinetics of the concepts of microscale intracrystalline tectonic overpressure and of macroscale local or regional tectonic overpressure, and for the hypotheses of subduction and obduction involving a more than 100 km vertical component of transport.

ALTHOUGH the rare mineral coesite, a high-pressure polymorph of SiO_2 , is well known in certain nodules in kimberlite pipes¹, in tektites², in craters formed by meteorite impacts or by TNT or nuclear explosions² and in very high pressure experimental syntheses³, it was completely unknown in the terrestrial crust until very recently, when it was discovered as a relic in SiO_2 inclusions in garnet, in an eclogite facies garnet-quartzite in the Alpine orogen of Italy^{4,5}. Here I describe the second known occurrence in a crustal environment (previously mentioned very briefly in refs 6 and 7); it differs significantly from that in Italy in the chain-silicate nature of the host mineral, the basic bulk-rock chemistry of the host rock, and its much greater age (≥ 400 Myr).

Sample

Rock sample C51 was collected from the complex of eclogite pods at Grytting in the Selje district in the northwesternmost part of the Western Gneiss Region (WGR) of Southern Norway. This locality⁸, best known for its beautiful layers of 'pegmatitic' coarse-grained eclogite containing magnesite $\pm$ orthopyroxene $\pm$ phlogopite⁹⁻¹¹, has recently become a protected site (sample C51 was collected with many others several years ago but has only recently been sectioned); the locality is mapped in ref. 10 and photographed in ref. 12.

The weakly-layered rock is predominantly composed of millimetric stoichiometric clinopyroxene ($\text{Jd}_{47}(\text{Di} + \text{Hd})_{45}(\text{En} + \text{Fs})_5\text{CaTs}_3$) and garnet ($\text{Pyr}_{33}\text{Alm}_{46}\text{Sp}_{11}\text{Gr}_{20}$) with $\sim 10\%$ quartz, $\sim 5\%$ dolomite, $\sim 1\%$ each of clinoamphibole, rutile and sulphides, and traces of zoisite/clinozoisite, ilmenite and zircon. Typical amphibolite-facies retrogression, forming plagioclase $\pm$ clinoamphibole microcrystalline to cryptocrystalline symplectites⁸⁻¹², is only locally developed, representing about 5% of the rock volume. The sample is thus assigned to the group of 'kyanite-lineage' dolomite $\pm$ zoisite eclogites defined by Lappin and Smith¹¹ which has been described in several localities in the Selje district; it came from a pod adjacent to the magnesite-eclogite pod described above.

Quartz occurs both in clear interstitial grains (200–2,000 μm long) and in inclusions (100–300 μm long) within clinopyroxene or garnet. Some of the inclusions in clinopyroxene contain 5–100- μm long relics of coesite surrounded by polycrystalline strained quartz (Fig. 1a, b) typical of that formed by the coesite $\rightarrow$ quartz transition in certain nodules in kimberlite¹ and in the Alpine garnet-quartzite^{4,5}. Other inclusions have either similar polycrystalline quartz without coesite relics (Fig. 1b), or a few larger crystals of unstrained quartz, usually 2–10 crystals per inclusion, which will be referred to as 'multicrystalline quartz' (Fig. 1c). Finally, some of the inclusions comprise only one clear quartz crystal (Fig. 1b). The whole sequence of events from minor transformation of coesite into quartz to complete quartz recrystallization is thus displayed in a single thin section (C51gt).

Coesite $\rightarrow$ quartz transition

Apart from its optical properties (a higher relief and a lower birefringence than of quartz) and its pure SiO_2 composition, the presence of coesite is indicated by the radial fractures in the clinopyroxene host crystal (Fig. 1a, b, c) caused by the $\sim 10\%$ volume increase of the coesite $\rightarrow$ quartz transition; similar radial fractures occur around the partially transformed coesite inclusions within garnet at the other localities^{1,4,5} but coesite has not previously been reported and confirmed within clinopyroxene. Only one garnet crystal in this thin section displays radial fractures around a polycrystalline quartz inclusion; although no coesite relics are visible, its earlier presence cannot be doubted. Of special interest is the development of concentric fractures in clinopyroxene around one inclusion (Fig. 1a).

The coesite relics in section C51gt were definitively identified with a Raman microprobe (RMP) which yielded spectra (such as Fig. 2) identical to those of other coesites confirmed by X-ray and electron diffraction⁷. As discussed by Boyer, Smith, Chopin and Lasnier (in preparation), all of the examined natural coesites from diverse localities produced spectral evidence for the presence of domains or crystallites of quartz within the coesite structure (see Fig. 2) which is believed to indicate incipient transformation to quartz.

It seems most likely that the interstitial quartz grains previously had the coesite structure but were more easily completely recrystallized because of the greater freedom of movement at the boundaries of the clinopyroxene and garnet grains (compare this with the interpretation of a pre-existing garnet-coesite in Italy⁴). However, a quartz-rutile vein reached within 3 cm of the thin section C51gt so that the interstitial quartz could be a later precipitate from a fluid phase.

There is no special textural evidence to suggest particular stress or shock conditions which could have generated coesite from quartz at a pressure less than that of its equilibrium stability field, as has been done experimentally in extreme strain-rate conditions¹³. Given the temperature (T) range of 700–850 $^{\circ}\text{C}$ as the minimum temperature of eclogite-facies equilibration of the carbonate-eclogites of the Selje district^{10,11}, the presence of coesite attests to a minimum pressure (P) of 27.5–29.5 kbar within the coesite crystals.

Given the characteristic textural features of pre-existing coesite (radial fractures emanating from polycrystalline or multicrystalline quartz inclusions, see above), it is quite probable that other coesite localities will soon be found in the Caledonides and in the Alps. Sample C254, a quartz-eclogite from the eclogite-ultrabasic-anorthosite-gneiss complex at Hessdalen⁹, Norway (50 km ENE of Grytting), displays such a texture (Fig. 1d) suggestive of the earlier presence of coesite, but this cannot be proved. The same clinopyroxene host crystal also displays well-developed orientated needles of quartz (Fig. 1d, structure confirmed by RMP⁷) exsolved from a pre-existing non-stoichiometric supersilicic pyroxene^{14,15}, a texture quite common

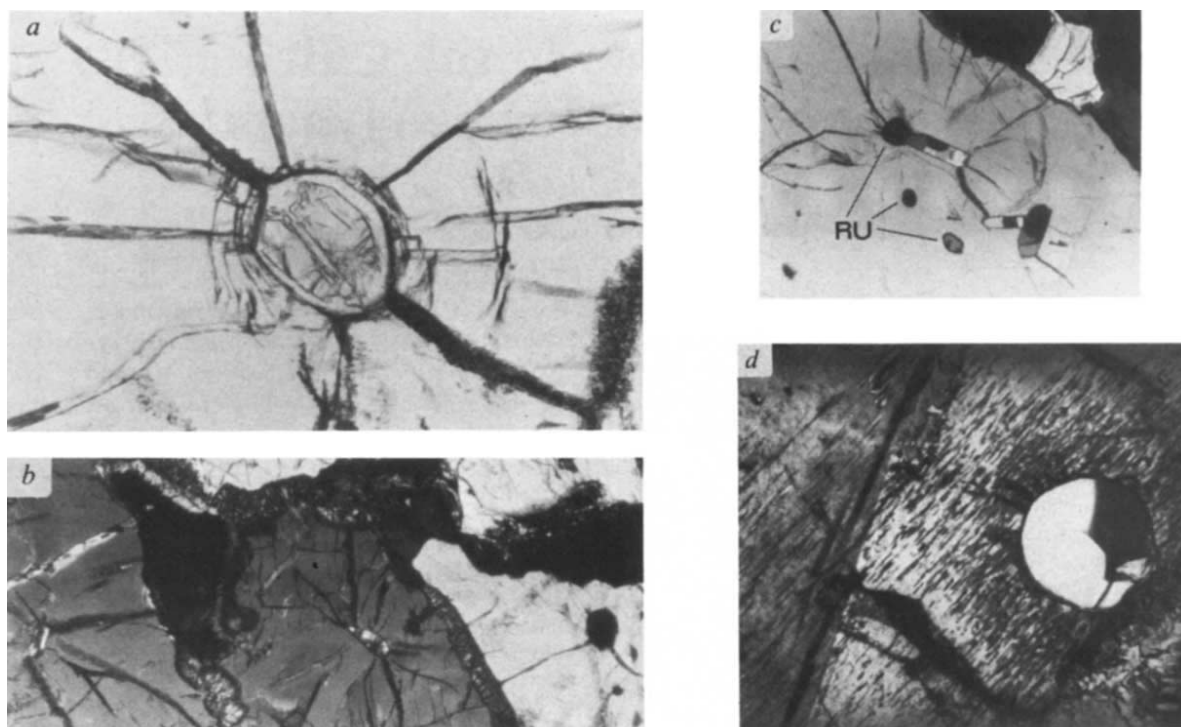


Fig. 1 SiO_2 inclusions in Norwegian eclogite clinopyroxenes: sample C51gt, Grytting (*a*, *b* and *c*); sample C254, Hessdalen (*d*); plane-polarized light, *a* and *c*; crossed polars, *b* and *d*. *a*, Coesite relic CS1. The inclusion ($140 \times 110 \mu\text{m}$) contains two large and several small relics of coesite (high relief) surrounded by polycrystalline quartz (white). Fractures in the host clinopyroxene radiate out from the inclusion towards the clinopyroxene grain boundaries. Concentric fractures are also visible. Minor late symplectitic alteration (dark fuzz) occurs locally. *b*, Three differing inclusions. Radiating fractures are evident in all three cases but are best developed around the central inclusion ($160 \mu\text{m}$ long) which is the only one still containing a coesite relic (CS5, $10 \times 10 \mu\text{m}$, indistinguishable with crossed polars but visible in plane-polarized light and confirmed by RMP). Polycrystalline quartz fills the central and left inclusions (variable white-grey-black birefringence) while a quartz monocrystal (dark grey) occupies the right inclusion. Also visible: garnet (black); dolomite (white, top right); another clinopyroxene (white, top centre); symplectite (dark fuzz). *c*, Two multicrystalline inclusions $\sim 300 \mu\text{m}$ long with straight parallel edges and comprising six or seven individual clear quartz crystals. One inclusion has a bent 'T'-shape (? twin relationship). Radial fractures are poorly developed. Other crystals (top right): garnet (black), clinopyroxene (white, dark-grey). RU, rutile inclusions. *d*, A round multicrystalline $200\text{-}\mu\text{m}$ diameter inclusion resembling those in Fig. 1*c*, comprising four quartz crystals (the black rectangle is a hole in the thin section). A series of short fractures radiate outwards. The clinopyroxene host (light-grey) and the clinopyroxene to the left (medium-grey) are full of regularly spaced parallel needles of quartz (confirmed by RMP) exsolved from pre-existing supersilicic pyroxene. Symplectitization occurs only locally (coarse lobate crystals, bottom right).

in quartz-eclogites in Norway and Greenland¹⁴. This sample may, in fact, provide an important clue to the origin of the coesite within clinopyroxene at Grytting. If at pressures below the coesite $\rightarrow$ quartz transition, excess SiO_2 in high-pressure supersilicic pyroxene is exsolved uniformly as orientated¹⁴ or non-orientated¹⁶ quartz, then perhaps at higher pressures excess SiO_2 can be exsolved as coesite which, for crystal-structural and/or crystal-growth reasons, tends to accumulate at a few isolated nucleation points.

Discussion

Crystal pressure is less easy to define and measure than rock of fluid pressure, especially when one approaches the Angstrom scale, but the pressure within an inclusion may exceed that within the rock due to different coefficients of compressibility and thermal expansion in the inclusion and host phases, as has been shown in crystal inclusions¹⁷ and in fluid inclusions (for example, higher pressure occurs in CO_2 inclusions within garnet than within quartz in some Norwegian eclogites¹⁸). Thus, micro-scale intracrystalline tectonic overpressure may have played some part in the preservation of the coesite inclusions. However, it is not thought that it was a significant effect in the creation of the coesite, especially in view of the facility for cleavage in clinopyroxene. It is thus deduced that $P \geq 30$ kbar affected the whole coesite-dolomite-clinopyroxene-garnet paragenesis.

If the eclogite C51gt was really subjected to a homogeneous minimum $P \sim 30$ kbar, and assuming the validity of the conventional translation of lithostatic pressure into depth, then ~ 100 km would be the minimum depth from which the rock must have been raised. If it had had a crustal protolith, which is quite likely, then it must first have been subducted to those depths along a rather high P/T gradient. A pressure as great as 30 kbar has previously been suggested for eclogites⁹⁻¹¹ and for ultrabases¹⁹ in Norway, but geophysicists and petrologists have not yet managed to integrate their efforts to understand the geodynamic problems involved^{20,21}; for example, many plate tectonic model cross-sections proposed by geophysicists still either ignore eclogites or treat only theoretically ideal eclogites which are rather different from natural ones which may contain several of the more than 40 different mineral species recorded in eclogites (see the table in ref. 12 and p. 330 of ref. 6). The mechanical problems of raising the dense eclogites now found within a continental setting is not confined to Norway as pressure estimates in eclogites from several other orogens have recently been revised upwards from 5–10 kbar to 15–20 kbar, thus corresponding to ~ 60 km depth, for example in Algeria (Hoggar²²), Austria (Tauern²³), France (Vendée²⁴), Germany (Münchberger²⁵) and Italy (Sesia-Lanzo²⁶ and Dora-Maira^{4,5} (> 20 kbar)).

A major development from the First International Eclogite Conference (1982) was the importance attached to the role of

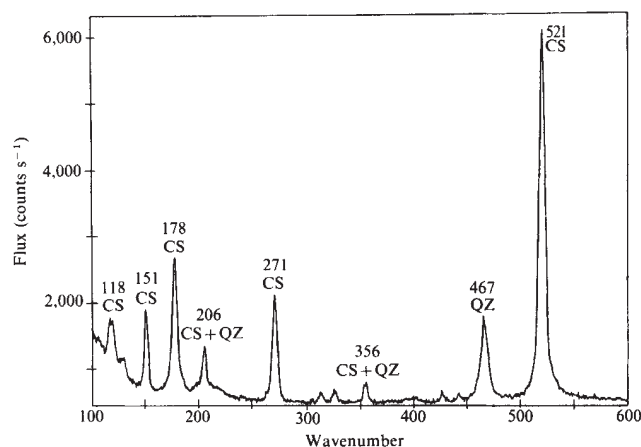


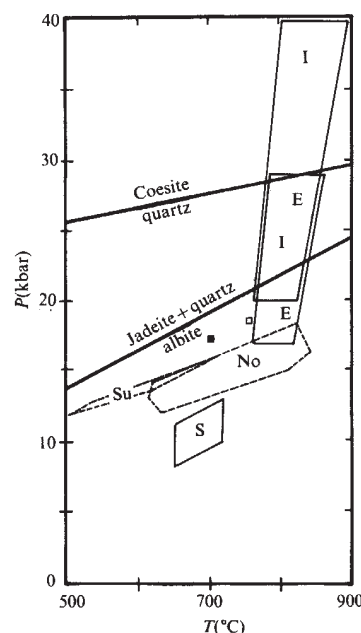
Fig. 2 An RMP spectrum from coesite relic CSI (Fig. 1a) exhibiting all the characteristic peaks of coesite (CS) and some of quartz (QZ). Laser: 400 mW at 514.5 nm; 300- μ m slits; 2-s counting at 0.5 wavenumber steps. For full details see Boyer, Smith, Chopin and Lasnier (in preparation).

deformation and shearing in the creation of eclogites (for example, in Algeria²², Italy²⁷, Norway^{28,29} and Spain³⁰), although there was considerable discussion over whether the key factor was a relative lack of reactivity in adjacent non-deformed rocks subjected to equal pressure, or an excess stress (some kind of tectonic overpressure) in the deformed rocks (see pp. 331–333 in ref. 6).

We are thus not constrained to assume that eclogites were in general subjected to homogeneous lithostatic pressure. The concept of tectonic overpressure is an old one but it has long been out of favour, although blueschist terranes, particularly in California³¹, New Caledonia³² and other circum-Pacific terranes, often provide geological evidence for such an overpressure as the estimated pressure cannot be accounted for by the maximum load of the presumed overlying material. Tectonic overpressures of more than a few hundred bars are not supposed to be possible in nature because of the limiting factor of the yield strength of rocks as measured experimentally; for example, Brace *et al.*³³ state that: "... tectonic overpressure depends upon rock strength. Obviously, if rocks undergoing metamorphism have little or no strength, then stress could not build up, no overpressure would exist...". However, the possible existence of small continuous pressure gradients, which would never approach the yield strength of the crystals and rocks concerned, does not seem to have been adequately disproved. Thus, for example, a pressure gradient of 1 bar mm⁻¹ (only 1 bar per average 1 mm-thick crystal) leads to 10 kbar per 10 m. Adding one or two orders of magnitude to allow for weaknesses in the boundaries of crystals and of rock units suggests that a 10 kbar pressure difference could exist for a certain time over only 100–1,000 m. Smith⁶ argued for the need to consider dynamic as well as static pressure; the former could perhaps be created by fluid expulsion from minerals, by shearing, or by heterogeneous horizontal movement during lithospheric interdigitation (for example, if two colliding slices (plates or nappes) become stuck together while the overlying and underlying slices slide above or below their counterparts).

It is thus suggested that dynamically stabilized non-lithostatic continuous pressure gradients could exist for long enough to allow pressure-sensitive mineral reactions to occur (many reactions in laboratory experiments need only a few hours or days). These inferred gradients could account for the generation of high-pressure rocks at lesser depths than if only homogeneous lithostatic pressure was applicable, thus alleviating some of the geodynamic problems of eclogite uplift. In other words, a dynamically produced overpressure might be sustained (by continually supplied stress) for some time before its dissipation—in

Fig. 3 *P*–*T* diagram showing some recent estimates from the WGR, Norway. Dashed polygons represent fields of specific estimates for orthopyroxene-free eclogites from Nordfjord (No) and Sunnfjord (Su); mean estimates for orthopyroxene-bearing eclogites (including Grytting): analysed Fe²⁺ (■); calculated Fe²⁺ (□) (all from refs 42–44). Solid quadrilaterals: uncertainty limits for the initial (I–I) amphibole-free eclogite-facies stage¹¹, early (E–E) amphibole-bearing eclogite-facies stage¹¹ and symplectitic (S) amphibolite-facies retrogression stage¹⁰ in eclogites with carbonate and/or orthopyroxene from the Selje district. All other published *P*–*T* estimates for gneisses or eclogites in the WGR lie below the Ab = Jd + Qz curve (from ref. 49) which represents the maximum pressure accepted by Griffin *et al.*⁴⁵ and others^{42–44,46}. Note that the position of the coesite $\rightarrow$ quartz curve³ supports the earlier ideas of a distinction in pressure between stages I and E^{9,10,50}.



a manner analogous to the sustenance of dynamically produced "overtemperature"⁶ around an igneous intrusion which causes contact metamorphism (due to inhomogeneous non-lithostatic temperature) before the eventual dissipation of the overtemperature.

This concept of a macroscale local or regional tectonic overpressure may seem startling, but dynamically stabilized non-lithostatic continuous pressure gradients occur every day in the laboratory (for example, a 40 kbar pressure difference between the centre and edges of a ~ 10 cm-, radius, traditional piston-cylinder press), so why should they not occur (even briefly) in nature?

The preservation of such dynamically formed high-pressure parageneses would seem to require simultaneous dissipation of pressure and temperature for example, by rapid tectonic uplift (possibly assisted by partially adiabatic processes), but this geodynamic/kinetic problem is also posed in the case of the alternative conventional homogeneous lithostatic pressure model.

Radial cracks around monocrystalline quartz inclusions within garnet, clinopyroxene or glaucophane in an eclogite from Shubino, USSR, and in garnet in an eclogite from Münchberg, Germany, were briefly reported many years ago³⁴. After discussing various possible effects of the compressibility or thermal expansion of quartz and the catalysis of coesite growth, those workers considered the retrograde transformation of pre-existing coesite into quartz to be the "most plausible" interpretation. Radial cracks around multicrystalline quartz inclusions in garnet in another eclogite (sample FTS4) from the Münchberger Gneiss Massif have been observed by Smith, Thomas and Franz (in preparation), thus corroborating Chesnokov and Popov's data³⁴, as does the inferred preservation in some Münchberger eclogites of pumpellyite at $T > 600$ °C (ref. 25) which would require an extremely high pressure.

Coesite is now known or indicated in four major Eurasian orogenic belts (Alps, Caledonides, Hercynides, Urals) and in xenoliths in the Siberian^{35,36} and South African¹ kimberlite provinces which may represent previously subducted crustal rocks^{6,37}. Coesite may thus be regarded as a mineral produced routinely during lithospheric plate collision, but only rarely does it manage to survive transport to the Earth's surface.

Norwegian eclogites

There are many facets of the international controversy over the origin of eclogites in Norway, such as the P - T conditions of metamorphism, geochronology, protolith type and the foreign tectonic introduction compared with *in situ* metamorphic origin. Some recent developments concerning eclogites in Norway are reviewed in refs 38–40, but the discovery of coesite has a profound impact on this controversy and merits a few comments here.

Equilibrium pressure in Norwegian eclogites, estimated by advocates of the *in situ* model, rose substantially during the 1970s from <10 kbar to <23 kbar (ref. 41) and has recently settled at ~20 kbar (refs 42–46). In particular, Griffin *et al.*⁴⁵ declared that " $P_{\max}$ is limited to 20 kbar since the albite=jadeite+quartz ($Ab=Jd+Qz$) transition was not exceeded". However, for the carbonate-bearing eclogites in the Selje district, very high P (~20 kbar or more) has been consistently maintained by advocates of the foreign model^{9–11,38,47,48} for the early (stage E) eclogite-facies equilibration [15–28 kbar (ref. 10); 17–29 kbar (ref. 11)]. Extremely high pressure was also tentatively deduced for the initial (stage I) eclogite-facies equilibrium [in the ranges 30–45 kbar (ref. 10); 20–40 kbar (ref. 11)], based on the low-Al content of orthopyroxene coexisting with garnet^{9,10,38} and on the coexistence of magnesite + diopside^{10,11} (Fig.

3), although Lappin and Smith^{10,11} emphasized that these estimates were open to doubt. The presence of coesite at Grytting now consolidates those extremely high pressure estimates and provides durable evidence that the $Ab=Jd+Qz$ reaction was exceeded (by at least 8 kbar at 800 °C, Fig. 3); it again underlines the importance of these carbonate-bearing eclogites.

Conclusions

The presence of coesite is believed to indicate $P \geq 30$ kbar in the whole dolomite-eclogite paragenesis at Grytting, rather than disequilibrium or substantial microscale intracrystalline tectonic overpressure. The hypotheses of uplift from ≥ 100 km depth (the homogeneous lithostatic pressure model) and of uplift from much lesser depths (the macroscale local or regional dynamically stabilized non-lithostatic tectonic overpressure model) are both retained, the latter being favoured slightly more than the former. In either case, current geodynamic models of lithospheric plate interaction need revising to take account of extraordinary vertical components of movement and/or of inhomogeneously distributed isobars.

I thank Mme H. Boyer and Drs C. Chopin and B. Lasnier for contributions to the RMP work, Dr G. Franz for constructive criticisms, Mr J. Barrandon for technical assistance and the Institut National d'Astronomie et de Geophysique for their grant (ATP 4542).

Received 8 February; accepted 18 June 1984.

- Smyth, J. R. & Hatton, C. J. *Earth planet. Sci. Lett.* **34**, 284–290 (1977).
- Stöfler, D. *Fortschr. Miner.* **49**, 50–113 (1972).
- Mirwald, P. W. & Massonne, H.-J. *Neues Jb. Miner. Mh* **469**–477 (1980).
- Chopin, C. *Terra Cognita* **3**, 183 (1983).
- Chopin, C. *Terra Cognita* **4**, 35–37 (1984).
- Smith, D. C. *Terra Cognita* **3**, 329–334 (1983).
- Boyer, H. & Smith, D. C. *Microbeam Analysis 1984* (San Francisco Press, in the press).
- Eskola, P. *Skr. Norske Vidensk.-Akad. Christiania Mat-Natur.* **K18**, 1–118 (1921).
- Smith, D. C. thesis, Aberdeen Univ. (1976).
- Lappin, M. A. & Smith, D. C. *J. Petrol.* **19**, 530–584 (1978).
- Lappin, M. A. & Smith, D. C. *Trans. R. Soc. Edinburgh Earth Sci.* **72**, 171–193 (1981).
- Smith, D. C., Kienast, J.-R., Kornprobst, J. & Lasnier, B. *Terra Cognita* **2**, 283–295 (1982).
- Green, H. W. *J. geophys. Res.* **77**, 2478–2482 (1972).
- Smith, D. C. & Cheeney, R. F. *26th Int. geol. Congr. Paris, Abstr.* **02.3.1**, 145 (1980).
- Smith, D. C. *Terra Cognita* **2**, 223 (1982).
- Smyth, J. R. *Am. Miner.* **65**, 1185–1191 (1980).
- Rosenfeld, J. L. & Chase, A. B. *Am. J. Sci.* **259**, 519–541 (1961).
- Touret, J., Smith, D. C. & Kechid, S.-A. K. *Terra Cognita* **2**, 318 (1982).
- Lappin, M. A. *J. Petrol.* **15**, 567–601 (1974).
- Smith, D. C. *Nature* **287**, 366–368 (1980).
- Franz, G. *Terra Cognita* **3**, 313–316 (1982).
- Sautter, V. *Terra Cognita* **2**, 321 (1982).
- Franz, G. & Spear, F. *Terra Cognita* **2**, 306–307 (1982).
- Godard, G. *Terra Cognita* **2**, 312 (1982).
- Franz, G., Thomas, S. & Smith, D. C. *Fortschr. Miner.* (in the press).
- Compagnoni, R. *Rend. Soc. Ital. Miner. Petrol.* **33**, 335–374 (1977).
- Kienast, J.-R. & Lombardo, B. *Terra Cognita* **2**, 303 (1982).
- Austrheim, H. & Griffin, W. L. *Terra Cognita* **2**, 315 (1982).
- Mork, M.-B. E. *Terra Cognita* **2**, 316 (1982).
- Morten, L., Bargossi, G. M., Martinez Martinez, J. M., Puga, E. & Diaz de Federico, A. *Terra Cognita* **2**, 320–321 (1982).
- Blake, M. C., Irwin, W. P. & Coleman, R. G. *Tectonophysics* **8**, 237–246 (1969).
- Brothers, R. N., Black, P. M. & Yokoyama, K. *Terra Cognita* **2**, 319 (1982).
- Brace, W. F., Ernst, W. G. & Kallberg, R. W. *Bull. geol. Soc. Am.* **81**, 1325–1338 (1970).
- Chesnokov, B. V. & Popov, V. A. *Doklady Akad. Nauk SSSR* **162**, 176–178 (1965).
- Sobolev, N. V. *et al. Doklady Akad. Nauk SSSR* **230**, 1442–1444 (in Russian) (1976).
- Lazko, Y. Y., Koptil, V. I., Serenko, V. P. & Tsepin, A. I. *Doklady Akad. Nauk SSSR* **258**, 138–142 (1983).
- Helmstaedt, H. & Gurney, J. J. *Terra Cognita* **2**, 272–273 (1982).
- Smith, D. C. *Terra Cognita* **2**, 304 (1982).
- Gebauer, D. *Terra Cognita* **3**, 320–323 (1983).
- Harte, B. *Terra Cognita* **3**, 317–319 (1983).
- Krogh, E. J. *Nature* **267**, 17–19 (1977).
- Griffin, W. L. *et al. Terra Cognita* **1**, 48–49 (1981).
- Griffin, W. L. & Mork, M.-B. E. *Uppsala Caledonide Symp. Excursion Guide B1*, 1–88 IGCP Proj. 27 (1981).
- Carswell, D. A., Krogh, E. J. & Griffin, W. L. *Terra Cognita* **1**, 39 (1981).
- Griffin, W. L., Mork, M.-B. E. & Austrheim, H. *Terra Cognita* **3**, 185 (1983).
- Cuthbert, S. J., Harvey, M. A. & Carswell, D. A. *J. metamorp. Geol.* **1**, 63–90 (1983).
- Smith, D. C. *26th int. geol. Congr. Paris, Abstr.* **01.3.3**, 92 (1980).
- Smith, D. C. *Terra Cognita* **1**, 73–74 (1981).
- Holland, T. J. B. *Am. Miner.* **65**, 129–134 (1980).
- Smith, D. C. *Norsk Geol. Tidsskr.* **51**, 141–147 (1971).

Comparative biochemical properties of normal and activated human *ras* p21 protein

John P. McGrath, Daniel J. Capon, David V. Goeddel & Arthur D. Levinson

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

Human Ha-ras1 cDNAs encoding normal and activated p21 polypeptides have been efficiently expressed in Escherichia coli and the biochemical activities associated with each polypeptide compared. In addition to the guanine nucleotide binding activity, normal p21 displays a GTPase activity which is selectively impaired by a mutation which activates its oncogenic potential.

DNA from cell lines derived from many human and rodent tumours contain altered *ras* genes capable of the tumorigenic and morphological transformation of NIH 3T3 murine fibroblasts^{1–5}. There are three known members of the human *ras* gene family, c-Ha-*ras*, c-Ki-*ras* and N-*ras*, the classification depending on the extent of homology with the oncogenes of Harvey (Ha) and Kirsten (Ki) sarcoma viruses and with the transforming gene first encountered in the SK-N-SH neuroblastoma line (see ref. 6 for a recent review). All members of the

ras gene family encode immunologically related proteins of molecular weight ~21,000 which have been termed p21s⁷. The cloning and characterization of viral *ras* genes⁷ and of normal and activated cellular Ha-*ras*, Ki-*ras* and N-*ras* genes have recently been described^{8–15}. The polypeptides encoded by these genes consist of 188–189 amino acids, diverge from one another most notably in their carboxy-terminal regions and are oncogenic when particular amino acid substitutions occur at position 12 or position 61.

The functional alterations responsible for the transforming activity of activated human *ras* proteins are unknown. Subcellular fractionation and electron microscopic evidence has indicated that both normal and oncogenic p21s are present at the inner surface of the plasma membrane^{16,17} and exhibit similar degrees of post-translational acylation¹⁸. Guanine nucleotide binding¹⁹, which is the only known biochemical activity of the human *ras* proteins, seems to be similarly unaffected by activating mutations¹⁸. In addition to its guanine nucleotide binding activity, v-Ha-*ras* p21 displays an autophosphorylating activity *in vitro*¹⁹ which is thought to account for the population of viral p21 which is phosphorylated on a threonine residue at position 59 *in vivo*²⁰. Cellular p21 is not phosphorylated²¹, presumably because it lacks the threonine residue at position 59 (refs 10, 22).

In an attempt to explore the biochemical basis for the transforming activity of activated human *ras* polypeptides, we have expressed both normal (glycine 12) and oncogenic (valine 12) versions of the human c-Ha-*ras* cDNA in *E. coli*, each with either the cellular (alanine) or viral (threonine) codon at position 59, introduced by site-directed mutagenesis. The nucleotide binding properties and autophosphorylation activities of these *ras* polypeptides have been examined. In addition to the ability to bind GTP specifically, and in some cases to autophosphorylate, we have detected a GTPase activity which extensively co-purifies with the bacterial p21 *ras* polypeptides. In contrast to the nucleotide binding potential, a mutation which activates the transforming potential of the polypeptide strongly impairs this GTPase activity.

Isolation of c-Ha-*ras*1 cDNA clones

To provide an enriched source of mRNA for the isolation of *ras*-specific cDNA clones required for bacterial expression of p21, we exploited vectors which place either the normal human c-Ha-*ras*1 gene (encoding a glycine at position 12) or the activated T24 bladder carcinoma homologue (encoding a valine at position 12) under the transcriptional control of the simian virus 40 (SV40) early promoter¹⁰. These plasmids contain an SV40 replication origin which permits their efficient propagation in COS-1 cells, a monkey cell line which is constitutive for SV40 T-antigen synthesis²³. After transfection of COS-1 cells the plasmids direct the synthesis of significant amounts of mRNA. We estimate that approximately 1% of the total poly(A)⁺ RNA contains *ras*-specific sequences by 72 h, representing an increase of 50-fold over the *ras* mRNA level found in T24 bladder carcinoma cells²⁴. mRNA isolated from COS cells transfected with the appropriate plasmids was used to prepare c-Ha-*ras* cDNA libraries (see Fig. 1 legend).

Nucleotide sequence analysis of such c-Ha-*ras*1 cDNAs revealed a continuous open reading frame of 567 nucleotides capable of encoding a polypeptide of 189 amino acids, with the normal and T24 bladder carcinoma-derived clones being distinguished by the alteration in the codon for amino acid position 12, as predicted from previous analysis of these genes^{10,22}. The three introns known to disrupt the c-Ha-*ras*1 coding sequences were correctly spliced during mRNA maturation. Each cDNA contained a 3'-untranslated region of 394 base pairs (bp) which was contiguous with the gene sequence, terminating with a poly(A) tract 16 nucleotides 3' of the hexanucleotide AGTAAA which may represent the polyadenylation signal sequence of the c-Ha-*ras* transcript (data not shown)^{10,22,25}.

Expression of c-Ha-*ras* p21 protein in *E. coli*

The approach taken to obtain efficient expression of p21 in *E. coli* involved the construction of plasmids which place the initiator methionine codon of each c-Ha-*ras*1 cDNA immediately adjacent to a strong bacterial promoter and ribosome binding site. A fragment encoding amino acid residues 6–110 of p21 *ras* was isolated from both normal and activated c-Ha-*ras*1 cDNA clones and ligated to a synthetic DNA fragment to restore the codons for the first five amino acids of p21 and to create an *Eco*RI cohesive terminus. The resulting fragment was joined to a DNA segment encoding the carboxy-terminal portion

Table 1 Specificity of GTPase activity and requirement for divalent cations

Unlabelled nucleotide	Relative dGTP hydrolysis
—	1.00
dATP	1.00
ATP	1.07
ADP	1.00
GTP	0.01
GDP	0.07
GMP	0.98
dCTP	0.96
CTP	0.81
CDP	0.93
dTTP	0.92
UTP	0.65
EDTA	0.02
Ca ²⁺	0.47
Mn ²⁺	0.84

Bacterially-derived p21 (Gly 12) was purified to ~95% homogeneity by extraction of the Nonidet P-40 (NP40)-insoluble bacterial fraction with urea (see Fig. 5), chromatography on Sephadex G75, dialysis against 50 mM Tris pH 7.5, 0.1 mM EDTA, and selective precipitation with 45% ammonium sulphate. GTPase reactions (0.1 ml) contained 0.1 µg p21, 50 mM Tris pH 7.5, 10 mM MgCl₂, 0.5 mM dithiothreitol (DTT), 1.3×10^{-8} M [α -³²P]-labelled dGTP, and, where indicated, unlabelled nucleotides at 1.3×10^{-6} M. To test divalent cation requirements, MgCl₂ was omitted from the reaction mixtures, replaced with 10 mM EDTA, 10 mM CaCl₂ or 1 mM MnCl₂, as indicated. The relative extent of dGTP hydrolysis was determined following resolution of the reaction products by chromatography on PEI cellulose plates as described in Fig. 5 legend.

of p21 and a portion of its 3'-untranslated region. The reconstructed gene was then inserted into a *trp* promoter expression plasmid²⁶, as illustrated in Fig. 1A.

Recombinant plasmids having the required structure, and encoding either glycine (pGAt_{trp}) or valine (pVAt_{trp}) at position 12, were identified and tested for the production of p21 by growing bacterial cells transformed with each plasmid in conditions of tryptophan depletion, which derepresses the *trp* promoter. Total cell extracts were prepared at various times of growth and proteins visualized following electrophoresis on SDS-polyacrylamide gels. As seen in Fig. 1B, cells transformed with pGAt_{trp} accumulate a prominent protein migrating with an apparent molecular weight of 23,000. A similar profile is observed in induced cells transformed with pVAt_{trp} (data not shown), while induced cultures of cells transformed with pBR322 do not express this polypeptide (compare lanes *j* and (-) in Fig. 1B) suggesting that it represents the translational product of the c-Ha-*ras* gene. Both proteins migrate with an apparent molecular weight slightly larger than that of mature p21, probably due to the inability of *E. coli* to carry out the post-translational modification (believed to represent an acylation reaction²⁷) which p21 undergoes in mammalian cells (our unpublished results). When grown in optimal induction conditions, this protein represents 5–10% of the total intracellular bacterial protein.

In vivo phosphorylation

One of the amino acid differences which distinguishes the c-Ha-*ras*1 gene product from its viral counterpart is the substitution of an alanine at position 59 for a threonine^{10,22}. In viral p21, this threonine residue serves as the substrate for phosphorylation *in vivo* and *in vitro*²⁰, whereas the cellular protein has not been found to be extensively phosphorylated²¹. Recently, hybrid cellular/viral *ras* genes have been constructed which encode p21 polypeptides. Those containing a threonine residue at position 59 are phosphorylated *in vivo* and apparently autophosphorylate *in vitro*²⁸. This suggests that human p21 is not phosphorylated because it lacks an appropriate phosphate acceptor site. To establish this directly, we introduced a specific mutation (GCC → ACC) into the two human p21 *ras* expression plasmids in order

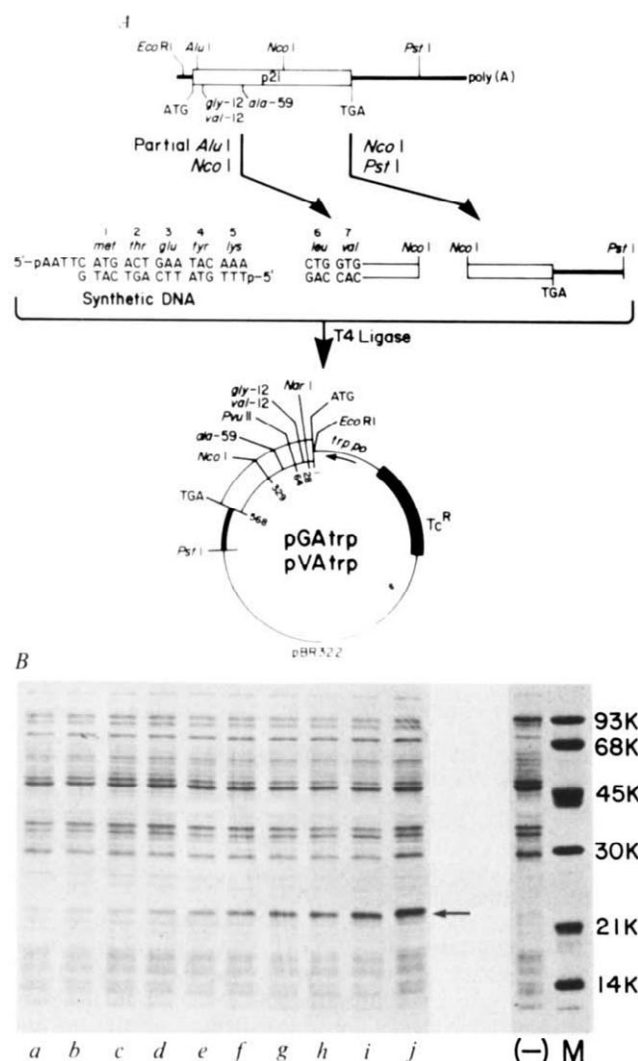


Fig. 1 **A**, Construction of Ha-ras p21 bacterial expression plasmids. c-Ha-ras1 cDNA clones were derived from COS-1 monkey cells transfected with SV40-Ha-ras chimaeric plasmids. The plasmids, pSVET24 (Val 12) and pSVEPBL (Gly 12), contain the human c-Ha-ras gene under the transcriptional control of the SV40 early promoter¹⁰. The coding region of p21 was assembled from a synthetic DNA duplex containing the first five amino acids and two fragments isolated from the cDNA clones, as shown. The synthetic DNA duplex incorporating an EcoRI site adjacent to the first five amino acids of p21 was joined to cDNA sequences encoding the remainder of p21 and inserted into a vector containing an *E. coli* *trp* promoter and ribosome binding site as shown. Coding and untranslated regions of the cDNA are indicated by an open box and solid line, respectively, as well as the positions of the codons specifying residues 12 and 59 (see text). **B**, Analysis of total cellular proteins in *E. coli* transformed with recombinant vectors. *E. coli* were transformed with pGAtrp and grown to the following A_{550} values in a fermenter: lane a, 8; b, 14; c, 20; d, 25; e, 30; f, 35; g, 38; h, 41; i, 44; j, 48. Lane (-) represents the profile from cultures transformed with pBR322 and grown to A_{550} = 48. Proteins were visualized by Coomassie blue staining following electrophoresis on 12% SDS polyacrylamide gels⁴⁹. Molecular weight markers (Bio-Rad) were electrophoresed in parallel (lane M). K represents 1,000 daltons.

Methods: **A**, COS-1 monkey cells were transfected with each plasmid separately by the DEAE-dextran method⁴¹; 3 days after transfection, total cytoplasmic RNA was prepared and poly(A)⁺RNA isolated by oligo(dT)-cellulose chromatography⁴². Double-stranded cDNA was prepared, using oligo(dT)₁₂₋₁₈ as the first strand primer^{43,44}, and cDNAs greater than 600 bp in size were extended with dCMP residues, using terminal deoxynucleotidyl transferase, annealed to PstI-digested pBR322 similarly tailed with dGMP residues and introduced into competent *E. coli*²⁶. The resulting tetracycline-resistant clones were screened by colony hybridization⁴⁵ using a radiolabelled probe prepared from the T24 oncogene *Bam*HI-*Sph*I fragment¹⁰. Full-length clones were identified by the presence of the EcoRI site introduced at the junction between the SV40 early promoter and c-Ha-ras1 gene. **B**, Total cell lysates from 3×10^7 cells were prepared for SDS-gel electrophoresis as described elsewhere⁴⁴.

to convert the alanine codon at position 59 to a threonine codon. This avoids the additional amino acid changes which typically accompany the design of cellular/viral hybrid genes. The strategy for site-specific *in vitro* mutagenesis is outlined in Fig. 2 legend. The appropriate mutations were then introduced into the original *trp* promoter expression vector (Fig. 1), yielding vectors pGTtrp (Gly 12, Thr 59) and pVTtrp (Val 12, Thr 59).

E. coli strains transformed with the four *ras* expression plasmids were analysed to determine whether the presence of a threonine residue at position 59 permits the human p21 to become phosphorylated; and if so, whether the amino acid 12 alterations associated with the activated protein influence the extent of phosphorylation. Bacterial cultures transformed with the four versions of p21 *ras* expression vectors were labelled with ³²P-orthophosphate for 1 h, after which time total cell extracts were analysed by SDS-gel electrophoresis and autoradiography. Figure 2 (left panel) illustrates the total polypeptide profile as visualized by Coomassie blue staining and confirms that comparable levels of the four *ras* polypeptides are present. It is noteworthy that each displays a unique electrophoretic mobility, consistent with previous descriptions of *ras* polypeptides which exhibit altered electrophoretic mobility as a consequence of single amino acid substitutions^{8,29}. As shown in the autoradiogram in Fig. 2 (right panel), human Ha-ras p21 is indeed phosphorylated when expressed in *E. coli*, but only when a threonine is present at position 59. The extent of phosphorylation in *E. coli* appears not to be affected by the activating mutation at codon 12, since similar levels of phosphate are incorporated into the glycine 12 protein as the valine 12 protein (compare lanes b and d). In addition, there is no

evidence to suggest that p21 catalyses the phosphorylation of any host proteins.

GTP binding properties

Predictions of the three-dimensional structure of the p21 molecule indicate that substitution of the glycine residue at amino acid position 12 might profoundly affect either the affinity or specificity of nucleotide binding^{30,31}. By using a protein blotting technique, we have directly assessed the relative affinity and specificity of nucleotide binding by normal and tumour-associated *ras* proteins. Unfractionated cell lysates from induced bacteria carrying the *ras* expression plasmids were resolved on SDS-polyacrylamide gels and electrophoretically transferred to nitrocellulose filters. The filters were then incubated in a binding buffer containing ³²P-labelled nucleotides, washed and autoradiographed. As illustrated in Fig. 3, lysates derived from *E. coli* transformed with c-Ha-ras expression plasmids contain a polypeptide which specifically binds [α -³²P]dGTP (or GTP, data not shown). No detectable binding was observed with either dCTP or dATP, properties characteristic of p21 *ras* isolated from mammalian cells¹⁹. The failure to observe a GTP binding component in bacteria not expressing this gene, and the exact co-migration of the guanine nucleotide binding components with the p21 proteins conclusively establishes the binding activity as a property intrinsic to the *ras* molecules.

The nucleotide binding experiments depicted in Fig. 3 indicate that there are no major alterations in the ability of p21 to bind guanine nucleotides which correlate with the activating lesion at positions 12, consistent with previous reports examining GTP binding activities associated with normal and activated p21s

immunoprecipitated from mammalian cells¹⁸. The specificity of nucleotide binding also appears unchanged between normal p21 and its tumorigenic variant. There are, however, significant differences in GTP binding affinities between the *ras* proteins with an alanine residue at amino acid 59 and those with a threonine at this position. As seen in Fig. 3, the threonine 59 mutants of p21 bind GTP less well than the alanine 59 versions of this protein. This mutation, however, has no apparent effect on the specificity of nucleotide binding.

Autophosphorylation of Thr 59-substituted p21

Both the normal and the activated versions of p21 engineered to contain a threonine residue at position 59 undergo phosphorylation *in vivo* as shown in Fig. 2, presumably because the presence of a suitable phosphoacceptor site enables the visualization of a cryptic autophosphorylating activity. This was directly confirmed by *in vitro* assays. Nitrocellulose filters bearing resolved whole-cell lysates of *ras*-expressing *E. coli* strains were incubated in the presence of [γ -³²P]dGTP (or GTP), washed as described above and autoradiographed. The profiles of polypeptides which associate with this labelled guanine nucleotide were similar to those observed with [α -³²P]dGTP (Fig. 4A). The filters were then washed for 2 h with 2% SDS at 60 °C, conditions expected to remove all non-covalently bound label. Figure 4B shows that these washing conditions remove all γ -labelled GTP from the p21 polypeptides containing alanine at position 59. In contrast, the threonine 59 mutants retain much of the ³²P label despite the vigorous washing manipulations, strongly indicating a covalent attachment of the γ -phosphate to this version of the polypeptide.

E. coli lysates containing the threonine 59 versions of the *ras* proteins display a p21 doublet after incubation with labelled dGTP (Figs 3, 4). The faster migrating species corresponds in size to unphosphorylated non-acylated p21, while the slower runs at the position expected for its phosphorylated derivative, pp21 (ref. 32). These experiments demonstrate that both p21 and pp21 can bind GTP and carry out autophosphorylation, and indicate that the isolated polypeptide is capable of exchanging the covalently bound phosphate. As neither species bearing an alanine at position 59 becomes phosphorylated, it seems that the threonine present at this position in the phosphorylated proteins serves as the phosphoacceptor site, although confirmation of this awaits direct analysis. With respect to their ability to catalyse this autophosphorylation reaction, a consistent difference is observed between the normal (Gly12) and activated (Val12) versions of the *ras* protein as measured by this assay. The activated polypeptide generally displays a 2–3-fold elevated level of *in vitro* autophosphorylating activity compared with the normal protein. In view of the observation that this enhanced *in vitro* activity is not reflected in *in vivo* experiments (Fig. 2), the significance of this observation is unclear.

p21-associated GTPase activity

The intrinsic ability of the *ras* proteins to function as phosphotransferases raised the possibility that they possess a coupled hydrolytic activity, as is often the case for phosphoryl group transferring enzymes³³. This was more specifically suggested by the functional and structural similarities between the family of *ras* proteins and several other cellular macromolecules such as transducin³⁴, elongation factor-Tu³⁵ and mitochondrial ATP-synthase³⁰, all of which have in common both purine nucleotide binding and nucleotide triphosphatase activities. To address this possibility, we assayed extracts of *E. coli* transformed with pBR322, pGAtpr and pVAtpr for the presence of a guanine nucleotide triphosphatase activity. As p21 associates primarily with the insoluble component of fractionated *E. coli*, this protein was enriched by isolating this fraction from induced bacteria. Polypeptides were solubilized in 8 M urea, and examined by SDS-polyacrylamide gel electrophoresis. This procedure enriches for two host proteins; in addition, significant amounts of a polypeptide corresponding precisely in electrophoretic

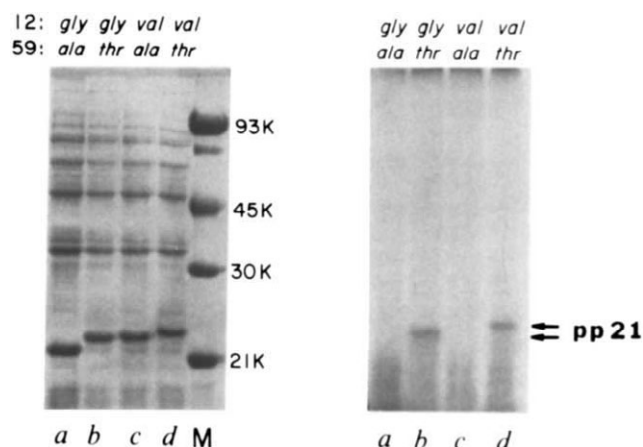


Fig. 2 Phosphorylation of human c-Ha-*ras* p21 in *E. coli*. Bacteria transformed with pGAtpr (lane a), pGTpr (lane b), pVAtpr (lane c) or pVTpr (lane d) (identical plasmids except for codons specifying amino acids 12 and 59 of p21; see text) were grown to $A_{550} = 1.0$ and labelled with ³²P for 1 h, after which time total cellular extracts were prepared for electrophoresis as described in Fig. 1 legend. Polypeptides were visualized by Coomassie blue staining (left panel), while total cellular phosphoproteins were detected after autoradiography of the Coomassie blue-stained lanes.

Methods: To convert the alanine 59 codon to threonine, *EcoRI*–*PstI* fragments containing this region from the p21 expression plasmids and encoding either glycine or valine at position 12 were cloned into M13mp10. *In vitro* site-specific mutagenesis was performed on the resulting single-stranded phage templates using a mismatched deoxyoligonucleotide primer (5'-CTGGATAC-CACCGGCCAGG) (threonine 59 codon emphasized) synthesized by the solid-phase triester method⁴⁶. Oligomers were phosphorylated using rATP and polynucleotide kinase and annealed to single-stranded mp10*ras* DNA. Annealings were carried out in 10 μ l buffer (10 mM Tris-HCl pH 7.5, 7 mM MgCl₂, 0.1 mM EDTA, 50 mM NaCl) containing 0.1 pmol of template DNA and 0.1 pmol phosphorylated oligomer and were incubated for 5 min at 55 °C. *In vitro* DNA synthesis was initiated by adding 40 μ l of 50 mM Tris-HCl pH 7.5, 7 mM MgCl₂, 0.1 mM EDTA, 1 mM ATP, 100 μ M each of dATP, dGTP, dTTP and dCTP, 5U of *E. coli* DNA polymerase I (large fragment; Boehringer Mannheim) and 400 U of T4 DNA ligase (New England Biolabs). After incubation for 3 h at 37 °C, reaction mixtures were extracted with phenol, DNA was precipitated with EtOH, and 2–4% of the DNA was used to transform competent *E. coli* strain JM103. Transformation mixtures (300 μ l) were spread onto bacterial plates (150 mm), yielding several hundred phage plaques. Plaques were screened for the presence of mutated DNA as described previously⁴⁷. Following identification of mutant progeny phage by hybridization with a radiolabelled mutagenesis primer, *EcoRI*–*PstI* fragments containing a threonine residue at position 59 and either a glycine or valine residue at position 12 were isolated from double-stranded phage replicative-form DNA (RF1) and cloned into a vector prepared from the original *trp* expression plasmids (Fig. 1). The structure surrounding the threonine 59 codon in the final expression plasmids was verified by DNA sequence analysis.

mobility to p21 are apparent (Fig. 5, left panel). When examined for the ability to hydrolyse GTP to GDP specifically, extracts prepared from bacteria expressing the normal human *ras* protein demonstrated a significant level of GTPase. This activity was markedly reduced in extracts containing equivalent amounts of oncogenic *ras* polypeptide, and virtually absent in extracts of cells transformed with pBR322 (Fig. 5, right panel). In conditions where the reaction was proceeding with linear kinetics, the normal (Gly 12) *ras* polypeptide hydrolysed GTP with a turnover eightfold greater than that of the activated (Val 12) polypeptide ($\sim 0.2 \text{ min}^{-1}$).

To establish that this GTP hydrolysis was directly mediated by p21, fractionated lysates from *E. coli* transformed with pBR322, pGAtpr and pVAtpr were electrophoresed in parallel on SDS-polyacrylamide gels, and individual gel fractions assayed for GTPase activity. Figure 6 illustrates that all the

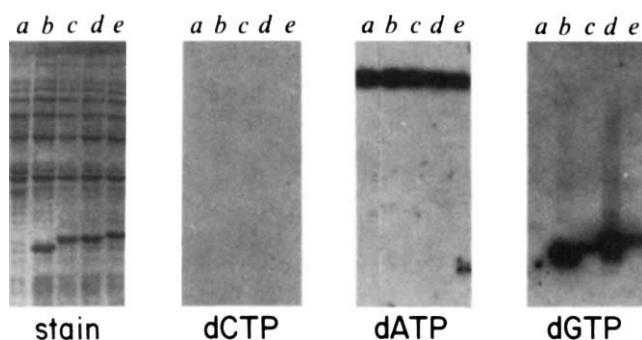


Fig. 3 Binding of [α - 32 P] nucleotides to p21 *ras*. *E. coli* transformed with pBR322 (lane a), pGAtrp (lane b), pGTtrp (lane c), pVAtrp (lane d) or pVTtrp (lane e) were solubilized in SDS sample buffer and electrophoresed on 12% SDS polyacrylamide gels. The resolved polypeptides were electrophoretically transferred to nitrocellulose filters⁴⁸ after first equilibrating the gel in transfer buffer (20mM NaPO₄, pH 8.0). The filters were incubated in 2 ml binding buffer (50 mM Tris pH 7.5, 5 mM MgCl₂, 0.1% NP40, 1 mg ml⁻¹ bovine serum albumin (BSA), [α - 32 P]-labelled nucleotides (50 μ Ci, 3,000 Ci mmol⁻¹; Amersham) as indicated for 1 h at 37 °C. The filters were then washed for 15 min with binding buffer lacking BSA and nucleotide to remove unbound label, dried and subjected to autoradiography.

GTPase activity detected in bacteria synthesizing normal (Gly 12) p21 co-migrated with this polypeptide. Fractions from extracts containing equivalent levels of activated (Val 12) p21 show a much reduced level of GTPase activity, which again co-migrates with the electrophoretically distinct activated species. These observations, together with the finding that extracts from bacteria transformed with pBR322 lack any detectable GTPase activity, clearly establish the GTPase as a property intrinsic to the *ras* polypeptides.

The specificity of p21 for guanine nucleotide binding has been well documented^{19,36}. To determine the specificity of p21 for nucleotide triphosphate hydrolysis, purified bacterially-derived normal (Gly 12) p21 was incubated with [α - 32 P]-labelled dGTP in the presence of a 100-fold excess of various unlabelled competitors, and the conversion of the labelled GTP to GDP was assessed. Table 1 illustrates that of the nucleotides tested, only GTP and GDP are able to competitively inhibit hydrolysis of the labelled dGTP substrate. GMP, ATP, ADP and various pyrimidine nucleotides compete poorly, if at all, paralleling the results obtained in nucleotide binding assays. The GTP hydrolysis activity displays an essential requirement for divalent cations, whereas nucleotide binding or autophosphorylation mediated by v-Ha-*ras* p21 exhibits no such requirement¹⁹.

Conclusions

To allow comparative biochemical analyses of the polypeptides encoded by normal and activated human *ras* genes, we have constructed plasmids which direct the synthesis of large quantities of these proteins in *E. coli*. While the expression of λ CI/v-*ras* gene fusions in bacteria has been described²⁷, we designed expression plasmids which programmed the synthesis of complete human *ras* polypeptides. This minimized the chance that fundamental differences between normal and activated *ras* proteins would be obscured by an inappropriate polypeptide structure induced by fused, heterologous amino acid sequences, particularly in view of suggestions that the extreme amino-terminal domain of p21 may participate directly in nucleotide binding^{30,31}. Significant quantities of an appropriately sized polypeptide are synthesized following induction of the *trp* promoter. Several lines of evidence establish this polypeptide as the direct translational product of the c-Ha-*ras* cDNA. First, the protein is specifically induced in cultures bearing the appropriate expression plasmid. Second, the electrophoretic mobility of the expressed *ras* polypeptides is affected by mis-

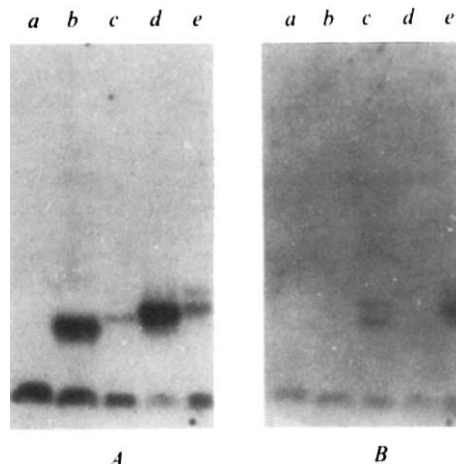


Fig. 4 Binding of [γ - 32 P]GTP to p21 *ras*. Lysates derived from bacteria transformed with pBR322 (lane a), pGAtrp (lane b), pGTtrp (lane c), pVAtrp (lane d) or pVTtrp (lane e) were resolved by SDS-gel electrophoresis, transferred to nitrocellulose filters and incubated in binding buffer as described in Fig. 3 legend. **A**, Binding of [γ - 32 P]GTP (50 μ Ci, 2,000 Ci mmol⁻¹; ICN). **B**, Profile of radioactivity remaining after washing the filter extensively with 2% SDS, 50 mM Tris pH 7.5 at 60 °C.

sense mutations engineered into the p21-expression plasmids. Third, direct amino-terminal sequence analysis of the first 12 amino acid residues of the bacterially expressed polypeptide reveals a complete correspondence with the predicted p21 sequence (J. Bell, R. Bott, K. McKeown and A. D. L., unpublished results).

Interest in the *ras* gene products has focused on two distinct regions within the molecule where mutational activation of the transforming potential has been demonstrated for both viral *ras* genes and activated cellular *ras* genes associated with tumours. To date, most naturally occurring Ha-*ras* activations have resulted from alterations of the codon specifying the glycine at residue 12, although changes at amino acid 61 can also activate the polypeptide. The amino-terminal sequence of p21 surrounding glycine 12 displays some homology with the nucleotide binding region in the β -subunit of mitochondrial and bacterial ATP synthase³⁰. In addition, the predicted structural configuration of this segment of the p21 molecule closely resembles that of the adenine dinucleotide binding unit of several enzymes³¹. These theoretical studies led to the suggestion that the amino-terminal region of the *ras* protein is involved in nucleotide binding, and indicated that changes in the glycine residue at position 12 might profoundly affect this function. We have tested these predictions by using a modified protein blotting technique which has been used previously in the study of a variety of protein-ligand interactions³⁸. By this procedure, we have demonstrated specific binding of labelled guanine nucleotides to human Ha-*ras* proteins synthesized in *E. coli*. However, comparative analysis of the normal gene product (Gly 12) and that encoded by the T24 bladder carcinoma oncogene (Val 12) reveals no obvious differences in either their affinity for GTP or their specificity for guanine nucleotides. A profound reduction in GTP binding was observed in *ras* polypeptides containing a threonine at position 59, but this was independent of an activating substitution at position 12. The significance of this observation awaits an evaluation of the biological properties of these polypeptides. It does, however, raise the possibility that Thr 59 may constitute a portion of the GTP binding domain, a notion consistent with its role as a phosphoacceptor site.

The observation that human p21 *ras* becomes phosphorylated both *in vivo* and *in vitro* when a threonine is present at position 59 clearly indicates that the human cellular p21 possesses this activity in a latent form. The demonstration that *E. coli* contains only extremely low levels of protein kinase activity^{39,40}, together with our findings that p21 represents the major phosphorylated species in bacteria expressing this gene product, lends strong

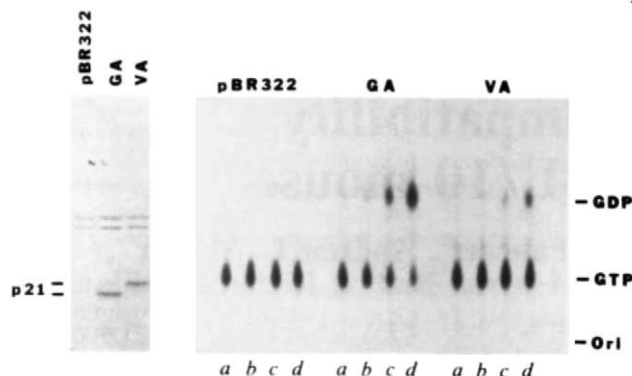


Fig. 5 Detection of GTPase activity in *E. coli* expressing human *ras* cDNAs. *E. coli* transformed with pBR322, pGAtpr or pVAtpr as indicated were grown to $A_{550} = 1.0$, and extracted with NP40 as described elsewhere⁴⁹. Sedimenting material was solubilized with 8M urea, and directly assayed for GTPase activity (right panel) by incubating 0.2–1.0 μ l (10–50 ng protein) with 0.1 ml [α -³²P]GTP (1 μ Ci, 3,000 Ci mmol⁻¹), 50 mM Tris pH 8.0, 5 mM MgCl₂, 0.1 mM DTT, 1 mM ATP for 10 min (lanes a), 20 min (lanes b), 60 min (lanes c) or 120 min (lanes d). The products of the reaction were resolved by chromatography on PEI cellulose plates (Merck) in 1 M LiCl. The positions of GDP and GTP were established by visualization of markers under UV light. Polypeptides extracted into urea were analysed by electrophoresis on 12% SDS polyacrylamide gels, stained with Coomassie blue (left panel).

support to the conclusion that pp21 observed *in vivo* results from an autophosphorylation event. Given this, our inability to discern any differences in the *in vivo* levels of phosphorylated p21 implies that the kinase activity is not significantly affected by mutations at position 12 which activate the transforming potential of the polypeptide. However, we did detect modest differences between the Gly 12 and Val 12 versions of p21 with respect to this activity when assayed *in vitro*, in agreement with the recent results of Gibbs *et al.*²⁸

The recent descriptions of the structural alterations responsible for the oncogenic activation of p21-*ras* have stimulated efforts aimed at determining the biochemical basis for the transforming activity of these mutant gene products. These attempts have failed to discern differences between normal and activated human *ras* polypeptides with regard to specificity or extent of nucleotide binding, acylation or subcellular localization, suggesting that some additional function may be critically associated with the activated state of the polypeptide. Thus, the detection of a p21-associated GTPase which is selectively impaired by an amino acid substitution that activates the transforming potential of the polypeptide, invites speculation that this activity may have a crucial role in the process by which *ras* polypeptides transform cells. How GTP hydrolysis might be

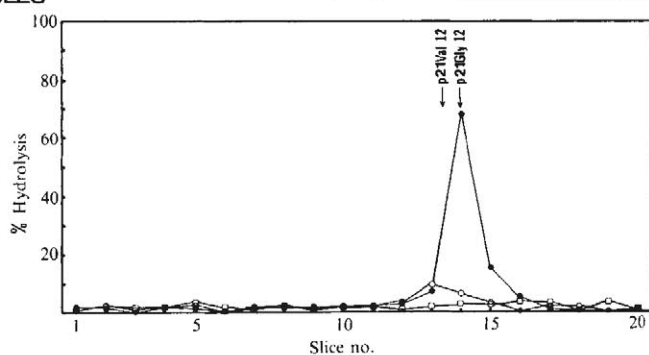


Fig. 6 GTPase activity co-migrates with the p21 polypeptide. Fractionated extracts (prepared as in Fig. 5) of *E. coli* transformed with pBR322 (□), pGAtpr (●) or pVAtpr (○) were electrophoresed on 12% SDS polyacrylamide gels. The gel was equilibrated in 0.2% NP40, 20 mM Tris pH 8.0, and individual fractions (each 0.6 mm) were assayed directly for GTPase activity by immersion of the gel slice in 0.1 ml buffer [50 mM Tris pH 7.5, 5 mM MgCl₂, 0.2% NP40, 1 mM ATP, [α -³²P]GTP (1 μ Ci, 3,000 Ci mmol⁻¹)]. After incubation for various times at 37 °C, aliquots were chromatographed as described in Fig. 5 legend. The per cent conversion of GTP to GDP was determined during the linear phase of the assay (up to 2 h) by Cherenkov counting of the isolated labelled components. The positions of p21 Gly-12 and p21 Val-12 were established by staining appropriate lanes run in parallel with Coomassie blue.

involved in the physiological function of either normal or activated *ras* polypeptides is unclear, but the demonstration of this activity increases still further the intriguing resemblance these proteins bear to the family of guanine nucleotide-binding proteins which regulate the activity of adenylate cyclase (G proteins)³⁴, as does the detection of significant stretches of homology between the α -subunit of bovine G proteins and members of the *ras* protein family (J. Hurley, J. Robinshaw, A. Gilman and M. Simon, personal communication). Such G proteins require bound GTP for activity, resulting in either the stimulation or inhibition of adenylate cyclase. Structural mutations that retard GTP hydrolysis might induce a similar phenotype as overproduction of normal α -subunit, the former by rendering the polypeptide in a constitutively active configuration, the latter by providing a supply of α -subunit free of the regulatory constraints of the β -subunit³⁴. If *ras* polypeptides are similarly involved in the transduction of information across the plasma membrane, this could rationalize the observations that either structural mutations in or overexpression of the normal *ras* gene⁵⁰ can induce the transformed phenotype.

We thank Mark Vassar for oligonucleotides, Ursula Edman for verification of plasmid sequences, and Mike Mumford for assistance with fermentation.

Received 16 July; accepted 23 July 1984.

- Shih, C., Shilo, B.-Z., Goldfarb, M. P., Dannenberg, A. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5714–5718 (1979).
- Cooper, G. M., Okenquist, S. & Silverman, L. *Nature* **284**, 418–421 (1980).
- Shih, C., Padhy, L. C., Murray, M. & Weinberg, R. A. *Nature* **290**, 261–264 (1981).
- Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1181–1184 (1981).
- Perucho, M. *et al.* *Cell* **27**, 467–476 (1981).
- Land, M., Parada, L. F. & Weinberg, R. A. *Science* **222**, 771–778 (1983).
- Ellis, R. W. *et al.* *Nature* **292**, 506–511 (1981).
- Tabin, C. J. *et al.* *Nature* **300**, 143–149 (1982).
- Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149–152 (1982).
- Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. *Nature* **302**, 33–37 (1983).
- Capon, D. J. *et al.* *Nature* **304**, 507–513 (1983).
- Shimizu, K., Goldfarb, M., Perucho, M. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 383–387 (1983).
- McGrath, J. P. *et al.* *Nature* **304**, 501–506 (1983).
- Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. *Cell* **34**, 581–586 (1983).
- Yuasa, Y. *et al.* *Nature* **303**, 775–779 (1983).
- Willingham, M. C., Pastan, I., Shih, T. Y. & Scolnick, E. M. *Cell* **19**, 1005–1014 (1980).
- Furth, M. E., Davis, L. J., Fleurdelys, B. & Scolnick, E. M. *J. Virol.* **43**, 294–304 (1982).
- Finkel, T., Der, C. J. & Cooper, G. M. *Cell* **37**, 151–158 (1984).
- Shih, T. Y., Papageorge, A. G., Stokes, P. E., Weeks, M. O. & Scolnick, E. M. *Nature* **287**, 686–691 (1980).
- Shih, T. Y. *et al.* *J. Virol.* **42**, 253–261 (1982).
- Papageorge, A., Lowy, D. & Scolnick, E. M. *J. Virol.* **44**, 509–519 (1982).
- Reddy, E. P. *Science* **220**, 1061–1063 (1983).
- Gluzman, Y. *Cell* **23**, 175–182 (1981).
- Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. *Nature* **296**, 404–409 (1982).

- Fasano, O., Taparowsky, E., Fiddes, J., Wigler, M. & Goldfarb, M. *J. molec. appl. Genet.* **2**, 173–180 (1983).
- Goeddel, D. V. *et al.* *Nature* **287**, 411–416 (1980).
- Sefton, B. M., Trowbridge, I. S., Cooper, J. A. & Scolnick, E. M. *Cell* **31**, 465–474 (1982).
- Gibbs, J. B., Ellis, R. W. & Scolnick, E. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2674–2678 (1984).
- Der, C. J. & Cooper, G. M. *Cell* **32**, 201–208 (1983).
- Gay, N. J. & Walker, J. E. *Nature* **301**, 262–264 (1983).
- Weirenga, R. K. & Hol, W. G. J. *Nature* **302**, 842–844 (1983).
- Shih, T. Y., Stokes, P. E., Smythers, G. W., Dhar, R. & Oroszlan, S. *J. biol. Chem.* **257**, 11767–11773 (1982).
- Morrison, J. F. & Heyde, E. A. *Rev. Biochem.* **41**, 29–54 (1972).
- Gilman, A. G. *Cell* **36**, 577–579 (1984).
- Leberman, R. & Egner, U. *EMBO J.* **3**, 339–341 (1984).
- Scolnick, E. M., Papageorge, A. G. & Shih, T. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5355–5359 (1979).
- Lautenberger, J. A., Ulsh, L., Shih, T. Y. & Papas, T. S. *Science* **221**, 858–860 (1983).
- Gershoni, J. M. & Palade, G. E. *Analyt. Biochem.* **131**, 1–15 (1983).
- Wang, J. & Koshland, D. E. *J. biol. Chem.* **253**, 7605–7608 (1978).
- Garnak, M. & Reeves, H. C. *J. biol. Chem.* **254**, 7915–7920 (1979).
- Sompayrac, L. M. & Danna, K. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7575–7578 (1981).
- Aviv, M. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
- Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253**, 2483–2495 (1978).
- Goeddel, D. V. *et al.* *Nature* **281**, 544–548 (1979).
- Grunstein, M. & Hogness, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961–3965 (1975).
- Crea, R. & Horn, T. *Nucleic Acids Res.* **8**, 2331–2348 (1980).
- Wallace, R. B., Schold, M., Johnson, M. J., Dembek, P. & Itakura, K. *Nucleic Acids Res.* **9**, 3647–3656 (1981).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
- McGrath, J. P. & Levinson, A. D. *Nature* **295**, 423–425 (1982).
- Chang, E., Furth, M., Scolnick, E. & Lowy, D. *Nature* **297**, 479–483 (1982).

Organization and evolution of the class I gene family in the major histocompatibility complex of the C57BL/10 mouse

Elisabeth H. Weiss^{*†}, Lynn Golden^{*†}, Karen Fahrner^{*}, Andrew L. Mellor^{*†}, James J. Devlin^{*}, Hilary Bullman[†], Harm Tiddens^{*}, Harriet Bud[†] & Richard A. Flavell^{*†}

^{*} Biogen Research Corporation, 14 Cambridge Center, Cambridge, Massachusetts 02142, USA

[†] Laboratory of Gene Structure and Expression, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

The major histocompatibility complex (MHC) encodes several classes of protein vital to the regulation of the immune response. We have isolated 26 class I genes that map to this region in the C57BL/10 mouse and linked these into three gene clusters. The number of genes differs from the number found in the BALB/c strain and comparison of the organization of the class I genes in these two strains shows conserved regions and polymorphic regions which probably result from deletions, insertions and translocations within the MHC.

CLASS I proteins of the murine major histocompatibility complex (MHC; *H-2*) consist of the classical transplantation antigens (K, D and L) found on virtually all cells (Fig. 1a) and appear to be involved in the presentation of foreign antigens, such as viral glycoproteins, to cytotoxic T lymphocytes. Class II antigens are believed to be involved in the presentation of foreign antigens to helper T lymphocytes and are found on the surface of a limited number of cell types including certain lymphocytes and macrophages¹.

Adjacent to the loci encoding these molecules on chromosome 17 is the *Tla* region which contains genes encoding the Qa and TL lymphoid differentiation antigens. These antigens, like the class I antigens, are integral membrane proteins with molecular weights of 40–45,000 and are expressed on the cell surface in association with β_2 -microglobulin². In addition, the genes encoding these antigens in both the *H-2* and *Tla* complexes are homologous; therefore these *Tla* complex genes are also referred to as class I genes.

One of the most intriguing aspects of the *H-2* class I antigens is their extreme polymorphism. More than 50 alleles at both the *H-2K* and *H-2D* loci have been observed (for a review see ref. 1). *H-2* haplotypes have been defined serologically and consist of a particular set of alleles found at each *H-2* locus, for example the C57BL/10 (B10) mouse is *H-2^b* at each *H-2* locus, thus expressing the *H-2K^b* and *H-2D^b* class I alleles. Many inbred mouse strains of different *H-2* haplotypes exist and a variety of recombinant lines have been generated from these strains, thereby providing a convenient system in which to study the evolution and function of this gene family.

Steinmetz *et al.*³ identified 36 class I genes in the BALB/c mouse (*H-2^d*), which they grouped into 13 clusters. Winoto *et al.*⁴ reported the location of these clusters within the MHC and corrected this gene number to 35. A comparison of the MHC of the B10 and BALB/c mice at the DNA level may help to explain the differential expression of certain class I genes (for example, *H-2L* and *TL*) in these mice.

Isolation of the class I genes

A total of 144 clones containing class I gene sequences was isolated from six cosmid libraries. We have arranged overlapping cosmids into three clusters containing a total of 26 different class I genes. One of these cosmid clusters spans the *H-2D* and *Qa2,3* regions; the other two clusters map to the *H-2K* and *TL* regions, respectively. The 5' to 3' orientations of the genes described here were determined by hybridization to the probes shown in Fig. 1b.

The gene clusters have been localized within the MHC using single or low-copy number probes from cosmids within each cluster. Each probe was hybridized to spleen DNA of B10

(*H-2^b*), DBA/2 (*H-2^d*) or AKR (*H-2^k*) mice that had been digested with a variety of restriction enzymes and Southern⁵-blotted. Probes that detected differences in the size of DNA restriction fragments among the inbred strains were then used to localize the polymorphic restriction fragment on the MHC map. This was done using congenic strains that have undergone recombination in the MHC. Serological studies have previously identified the haplotype that donated the DNA at each locus in these strains. Thus, an appropriate panel of recombinant strains allows the correlation of a particular haplotype of DNA at a given locus with the appearance of a particular polymorphic restriction fragment, thereby mapping the probe, and hence the gene cluster from which the probe was isolated, to that region.

We have used a nomenclature for the class I genes which reflects the genetic location of the gene cluster and the position of a gene within the cluster. Where a gene product has been identified, we use the established nomenclature. Thus, the two genes in the *H-2K* region are termed *K1* and *H-2K^b*; the 10 genes in the *Qa2,3* region are termed *Q1* to *Q10*, and the 13 genes in the *TL* region are termed *T1* to *T13*.

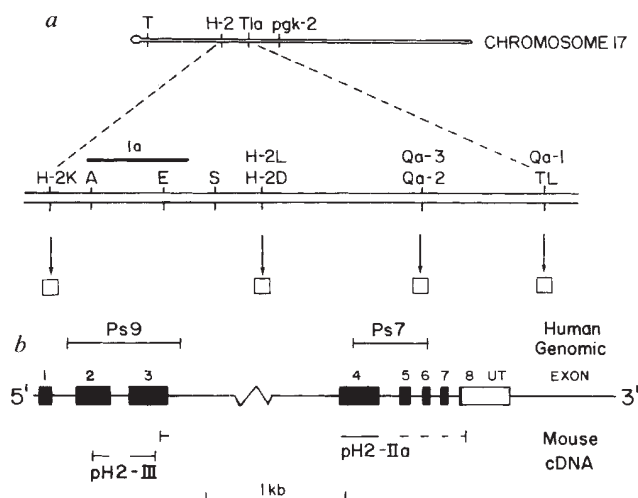


Fig. 1 a, Genetic map of the murine MHC. The centromere of chromosome 17 is located to the left. Class I molecules (□) are expressed from at least four separate loci as shown by the arrows below the line. The Ia region encodes the class II antigens. The S region encodes components of the complement system. b, Intron-exon organization of an MHC class I gene as described elsewhere¹⁸. Exons are shown as thick boxes (numbered 1–8). UT refers to the 3' untranslated region. Relative positions of human genomic²¹ and mouse cDNA²² clones used in the initial stages of this study are shown.

Table 1 Summary of the mapping of restriction fragments

Probe:	5' Flanking <i>K</i> region										3' Flanking <i>K</i> region					<i>D^b</i> gene mapping probe	<i>Q1</i> gene mapping probe	<i>TL</i> mapping probe
Restriction enzyme:	<i>Hind</i> III										<i>Bgl</i> II					<i>Bgl</i> II	<i>Hind</i> III	<i>Pst</i> I
Size (kb):	6.6	4.4	3.4	3.0	2.7	2.4	9.0	8.5	6.0	4.8	3.5	2.3	2.0	1.8	1.5	2.75	4.7	2.0
Gene:	<i>K1</i>	<i>K^b</i>	<i>Q6, Q8</i>	<i>Q10</i>	<i>Q7</i>	<i>Q5</i>	<i>Q2</i>	<i>D^b</i>	<i>T1</i>	<i>Q5, Q7</i>	<i>Q1</i>	<i>K^b</i>	<i>Q6, Q8</i>	<i>K1</i>	<i>T5</i>	<i>D^b</i>	<i>Q1</i>	<i>T1</i>
Strain																		
B10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
A	-	+	-	+	-	+	+	+	+	+	+	-	+	-	+	-(2.95)	+	+
DBA/2	-	+	-	+	-	+	+	+	+	+	+	-	+	-	+	-(2.95)	+	+
AKR	-	+	-	+	-	-	-	-	+	-	+	-	+	-	+	-(2.95)	-(3.3)	-(3.7)
B6.K1	+	+	+	+	-	-	-	+	+	-	+	+	+	+	+	+	-(3.3)	-(3.7)
B6.K2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-(3.7)
D2.R107	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-(2.95)	+	+
D2.GD	-	+	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+
B10.A(5R)	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
B10.A(2R)	-	+	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+
A.AL								+	+	+	+	+	-	+	-	+	+	+
C3H.OH								-	-	+	-	+	-	+	-	+	+	+
B10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Location	<i>H-2K</i>	<i>H-2K</i>	<i>Qa2,3</i>			<i>Qa2,3</i>	<i>H-2D</i>			<i>Qa2,3</i>	<i>H-2K</i>		<i>H-2K</i>		<i>H-2D</i>		<i>Qa2,3</i>	<i>TL</i>

The 5' flanking *K*-region probe is the 0.6-kb *Bam*HI fragment 2.6-kb 5' to the *H-2K^b* gene (Fig. 2a). The 3' flanking *K*-region probe is the 1.8-kb *Bgl*II fragment of H24 as indicated in Fig. 2a. As a *D^b* gene mapping probe, we isolated a 2.75-kb *Bgl*II fragment shown in Fig. 2b. The *Q1* mapping probe is a 4-kb *Pvu*II fragment, derived from the 7.2-kb *Kpn*I fragment of B1.24 indicated in Fig. 2b. To localize gene *T1* we used a 2.0-kb *Eco*RI fragment of H11 shown in Fig. 2c. + Indicates that the mapping probe detects a fragment of the same size in the B10 DNA and in the DNA of the respective haplotype; - indicates the absence of a band of the same size. If a B10 is absent and its corresponding fragment in another haplotype is known, its size is given in parentheses.

The *H-2K* region

This region is defined by 13 cosmid clones (five of which are shown in Fig. 2a) spanning 95 kilobases (kb) and contains two class I genes. The two genes are spaced about 15 kb apart and are arranged in a head-to-tail configuration with no other class I genes for at least 30 kb on either side.

We have established the genetic location of this cluster by showing that one of the genes in this region is the *H-2K^b* gene, both by DNA sequence analysis⁶ and by showing that it directs the expression of *H-2K^b* antigens in L-cells transformed by cosmids or subclones containing this gene⁷. This result has been confirmed by polymorphic restriction site mapping using a 1.8-kb *Bgl*II fragment (the 3' flanking *K*-region probe) isolated from cosmid H24 which contains exons 6 to 8 and 3' flanking sequences of gene *K1* (Figs 2a, 3a, Table 1). When used as a hybridization probe on genomic blots of B10 DNA digested with *Bgl*II, this fragment hybridizes strongly with nine fragments (Fig. 3a, Table 1). As shown by the hybridization patterns in Fig. 3a and the genetic maps of the recombinant inbred strains in Fig. 3e, this fragment maps to the *H-2K* side of the *I-E* locus (see, for example, B10A.5R). Of the remaining eight DNA fragments detected, one flanks the *H-2K^b* gene, one flanks the *H-2D^b* gene and the remainder are found flanking class I genes in the *Tla* complex.

The *H-2D* and *Qa2,3* region

This region is defined by 67 cosmid clones (21 of which are shown in Fig. 2b) spanning 320 kb and contains 11 class I genes. One of these genes was identified as the *H-2D^b* gene by transformation of L-cells with cosmids or subclones containing this gene and analysis with *H-2D^b*-specific antibodies⁸ and cytotoxic T cells⁹. The location of this gene was confirmed using two mapping probes. A 2.7-kb *Bgl*II fragment isolated from the region shown in Fig. 2b maps to the *H-2D* region (Table 1). In addition, the 3' *K*-region flanking probe hybridizes to a polymorphic 8.5-kb *Bgl*II fragment starting in this gene, which also maps to the *H-2D* region (Fig. 3a, Table 1).

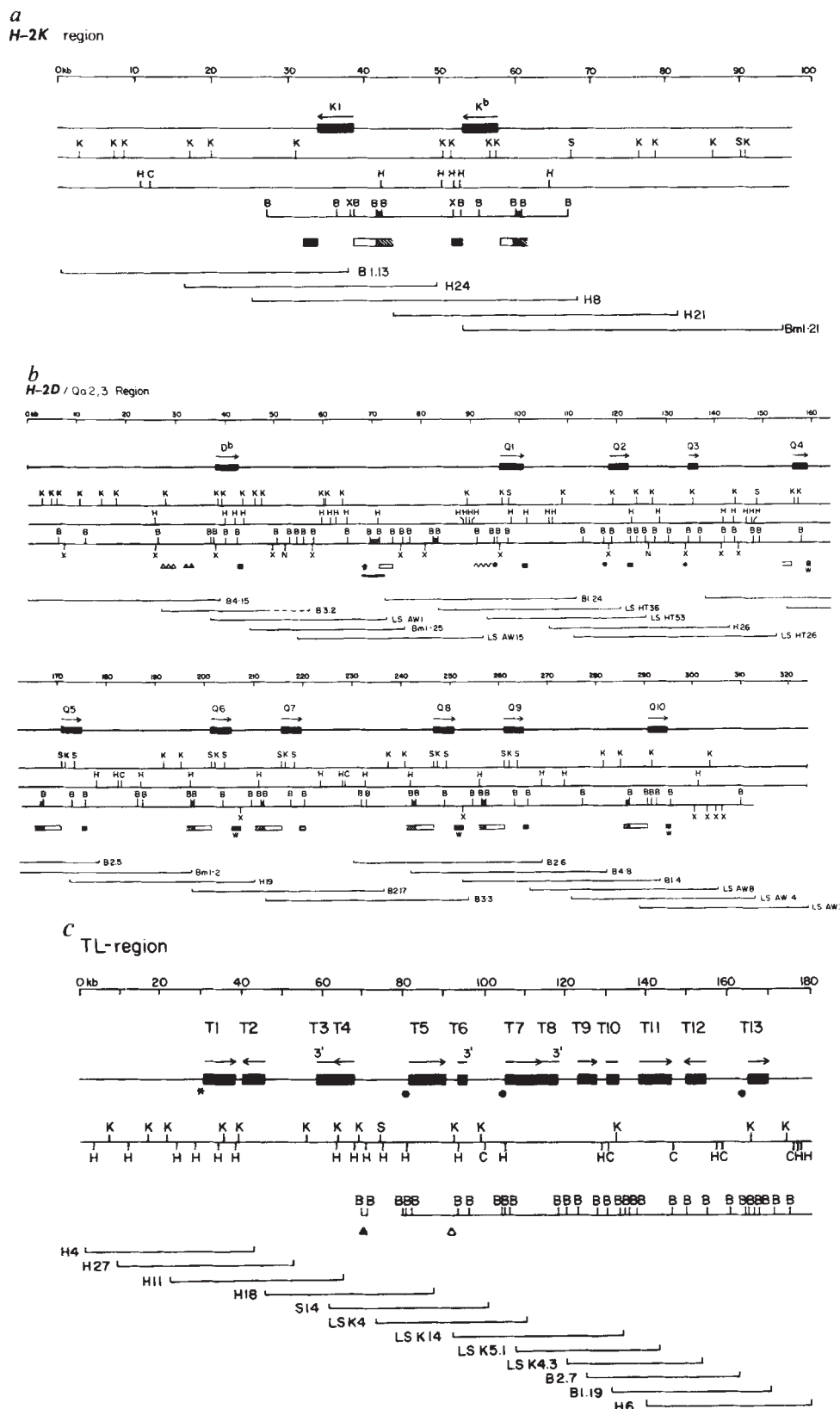
The other 10 genes in this cosmid cluster are located in the *Qa2,3* region. The *Q1* gene was mapped to the *Qa2,3* region

using the 4-kb *Pvu*II fragment just 5' of the gene (Fig. 2b). This probe hybridizes to a 4.5-kb *Hind*III fragment in B10 and B6.K2 DNA, but to a 3.3-kb fragment in AKR and B6.K1 DNA (Fig. 3b, Table 1). Thus, the site of recombination between the *Qa2,3* region and the *H-2D* region in the B6.K1 mouse must be in the 40-kb DNA between the 8.5-kb *Bgl*II fragment of the *H-2D^b* gene and 4-kb *Pvu*II fragment adjacent to the *Q1* gene (see Fig. 2b). Genes *Q2*, *Q5*, *Q7* and *Q9* were mapped to the *Qa2,3* region independently using the 3' flanking *K*-region probe which hybridizes to polymorphic *Bgl*II fragments (Fig. 3a, Table 1).

Additional information on the location and relationships among the class I genes was obtained using a 0.64-kb *Bam*HI fragment (the 5' flanking *K*-region probe) isolated from a position 2.6-kb to the 5' side of the *H-2K^b* gene (Fig. 2a). This probe hybridizes to six fragments of *Hind*III-digested B10 DNA (Fig. 3c). Two polymorphic *Hind*III fragments (6.6 and 3.4-kb; Fig. 3c, Table 1) map to the *H-2K* region of the B10 mouse (Fig. 2a). The other polymorphic fragments (2.7 and 2.4-kb) map to the *Qa2,3* region (Fig. 3c, Table 1) and are found in the 5' flanking regions of genes *Q7* and *Q9* (2.7-kb fragment) and gene *Q5* (2.4-kb fragment). Two non-polymorphic *Hind*III fragments are also contained in cosmids from this cluster: the 3-kb fragment is found in the 5' flanking region of the *Q6*, *Q8* and *Q10* genes; the 4.4-kb fragment is located between the *H-2D^b* and *Q1* genes (Fig. 2b). Thus, except for this last case, the 5' flanking *K*-region probe always hybridizes to regions which are 2.6-3.8-kb to the 5' side of class I genes in the *H-2K* and *Qa2,3* regions.

We were able to link the *H-2D* and *Qa2,3* regions after screening cosmid libraries with the 5' flanking *K*-region probe and isolating cosmids containing the 4.4-kb *Hind*III fragment described above. This fragment does not flank a class I gene; however, the 2.3-kb of DNA adjacent to it is homologous to the DNA sequences immediately flanking the *H-2K^b* gene on its 5' side; this indicates that a gene may once have been present at this site. Inspection of the restriction map of cosmid LS-AW1, which contains the *H-2D^b* gene, showed significant overlap with cosmid Bm1-25. Moreover, cosmid LS-AW15 overlaps with cosmid B1.24, which contains gene *Q1* (Fig. 2b); this overlap was confirmed by another series of 12 cosmids (not shown).

Fig. 2 Restriction maps of the three cloned regions. The length of each region (kb) is shown on the top line. The following sites are indicated: B, *Bam*HI; C, *Cl*al; H, *Hpa*I; K, *Kpn*I; N, *Nru*I; S, *Sal*I; X, *Xho*I. Class I genes are shown as solid boxes with arrows indicating the direction of transcription, 5'→3'. A representative set of overlapping cosmids defining each region is given below the restriction map. **a**, *H-2K* region. **b**, Regions hybridizing to the 3' *K*-region flanking probe (1.8-kb *Bgl*II fragment) isolated from cosmid H24 at map position 32–34. Regions hybridizing to four different 5' flanking region probes are indicated: **■** on the line showing the *Bam*HI restriction sites indicates *Bam*HI fragments that hybridize to the 5' *K*-region flanking probe (0.64-kb *Bam*HI fragment 5' to the *H-2K^b* gene). **▨**, Regions which cross-hybridize to the 1.8-kb *Bam*HI fragment isolated from cosmids spanning the region between the *Q1* and *H-2D^b* genes (see **b** below). **□**, Regions which hybridize to two probes spanning the 2.45-kb region immediately 5' to the coding sequence of the *H-2K^b* gene (1.75-kb *Pst*I fragment and 0.7-kb *Pst*I–*Kpn*I fragment; see ref. 6). *Bam*HI sites are shown only for the region between 25 and 68-kb. **b**, *H-2D/Qa2,3* region. The *Q3* gene contains only exons 1–3 of a class I gene. Most of the overlapping cosmid clones shown were isolated by screening libraries with class I gene probes. In some cases overlaps were confirmed by screening cosmid libraries with low-copy number gene flanking probes or walking probes. The region between the *Q1* and *H-2D^b* genes is defined by cosmids which were isolated using either the 5' *K*-region flanking probe (Bm1-25) or the *Qa* walking probe (LS-AW1 and LS-AW15). The first probe hybridizes to a 4.4-kb *Hind*III fragment (see text and Fig. 3c) shown as the solid bar at 70-kb on the map, and the second probe (1.8-kb *Bam*HI fragment indicated by **▨** on the *Bam*HI restriction site line) was isolated from this same region. Note that this probe hybridizes twice in this region (at 70 and 82-kb on the map) and to the 5' flanking region of genes *Q5* to *Q10*, *K1* and *H-2K^b*. Other symbols: $\Delta\Delta\Delta$, the location of the 2.75-kb *Bgl*II fragment used as an *H-2D* region mapping probe; $\blacktriangle\blacktriangle$, the region which cross-hybridizes to a DNA fragment isolated from the region flanking the *H-2L^d* gene⁴; $\sim$, the 4.0-kb *Pvu*II fragment used as *Qa2,3* region mapping probe; $\bullet$, regions which cross-hybridize to a 0.5-kb *Bam*HI fragment isolated from the 5' flanking region of the *T7* gene (c). Note that this probe also hybridizes to the previous (4.0-kb *Pvu*II) probe. For other symbols see **a**. W indicates weak hybridization to the 3' flanking *K*-region probe. The dotted line at the right end of cosmid B3.2 indicates that this segment is rearranged. **c**, *TL* region. Cosmids defining this region were isolated by screening cosmid libraries with class I gene probes or using low-copy probes. $\blacktriangle$, A 0.7-kb *Bam*HI probe isolated from cosmid H18 which also hybridizes to a region between genes *T5* and *T6*. *, Location of the 2.0-kb *Eco*RI fragment obtained from clone H11 which was used as a *TL* region mapping probe. $\bullet$, A probe isolated from the 5' flanking region of gene *T7* (0.54-kb *Bam*HI fragment) cross-hybridizes to the 5' flanking regions of genes *T5* and *T13* as well as to genes in the *Qa2,3* region (Fig. 2b). The orientation of gene *T10* has not been determined. Genes *T3*, *T6* and *T8* may be 3' gene fragments since they do not have adjacent sequences which hybridize to 5' class I gene probes. Construction of cosmid libraries: libraries were screened and clones picked using procedures developed previously²³ incorporating the modifications of Ish-Horowitz²⁴ and using cosmid vectors constructed in our laboratory²⁵. Five libraries were used. Cosmids with prefix H are from a library⁷ made with B10 spleen DNA and vector pOF1; library B, from B10 spleen DNA using vector pTM; library LS, from B10 liver DNA using vector pTCF; library S was made with B10 spleen DNA using pTCF as vector. Finally, one library (Bm1) was made using liver DNA from the B6 mouse strain which carries the *H-2K^{bml}* mutation²⁶ and vector pTM.



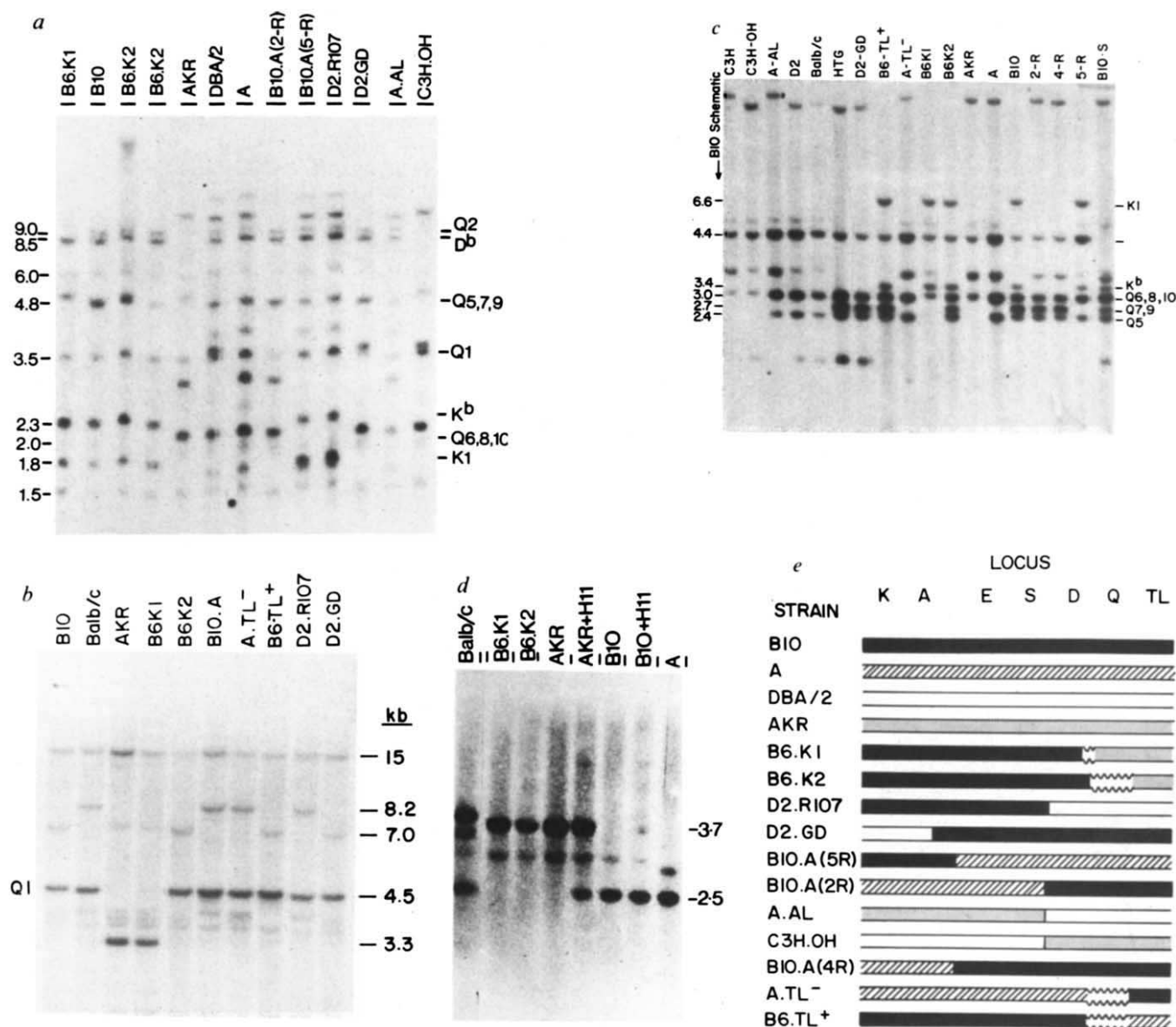
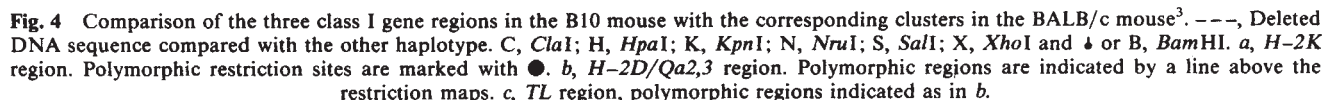


Fig. 3 Southern⁵ blot analysis of different digests of murine DNA with single or low-copy probes from the *H-2K*, *Qa2,3* and *TL* regions of the B10 mouse. Southern blots were hybridized at 65 °C in 3 × SSC, 10 × Denhardt's solution containing 0.1% SDS, 0.1% sodium pyrophosphate, 10% dextran sulphate^{7,23,27}; 20 µg ml⁻¹ denatured murine DNA was added to compete with the hybridization of repeated sequences. The filters were given a final wash with 0.1 × SSC, 0.1% SDS at 65 °C. DNA fragments were isolated for subcloning and used as hybridization probes by the method of Weislander²⁸. **a**, *Bgl*III digest, probed with the 3' flanking *K*-region probe (1.8-kb *Bgl*III fragment from clone H24; Fig. 2a). The sizes (kb) of the nine strongest B10 bands are indicated on the left. The class I genes most closely linked to each of the nine bands are given on the right. The weakly hybridizing 6.0- and 1.5-kb *Bgl*III fragments lie in the 3' ends of the *TL* and *T5* genes, respectively. The 4.5-kb *Hind*III fragment lies 5' to the *Q1* gene. **b**, *Hind*III digest, probed with the *Q1* mapping probe (4-kb *Pvu*II fragment, 5' to the *Q1* gene; see Fig. 2b). The sizes of the bands (kb) are indicated on the right. **c**, *Hind*III digest, probed with the 5' flanking *K*-region probe (0.64-kb *Bam*HI fragment; Fig. 2a). The sizes of the six B10 bands are given (kb) on the left and the closest linked gene(s) on the right. The non-polymorphic 4.4-kb *Hind*III fragment is located between the *H-2D^b* and *Q1* genes (Fig. 2b). **d**, *Pst*I digest, probed with the 2-kb *Eco*RI mapping probe from cosmid H11 (Fig. 2c). The sizes of the strongest bands (kb) are given on the right. Two tracks (+H11) contain H11 cosmid DNA digested with *Pst*I mixed with AKR- or B10-restricted DNA. **e**, Genetic maps of mouse strains used to localize cosmids to MHC loci. HTG (not shown) is *H-2K^d*, *I-A^d* and *H-2D^b*. ≈, DNA in these regions could be derived from B6 or A.

The *H-2D^b* gene is in the same 5' to 3' orientation as the genes in the *Qa2,3* region (Fig. 2b).

The *Qa2,3* region map suggests that this region has evolved by a series of sequential (or possibly simultaneous) gene duplications. Related class I genes can be identified by the hybridization pattern of four different 5' flanking probes and a single 3' flanking *K*-region probe (Fig. 2a, b). Each probe hybridizes to the respective flanking sequences of genes *Q5* to *Q10* (Fig. 2b) as well as genes *K1* and *H-2K^b* (Fig. 2a). The hybridization of

two of the 5' flanking probes to gene *Q4* suggests a more distant similarity between gene *Q4* and genes *Q5* to *Q10*. Furthermore, the restriction maps of DNA flanking the *Q5*, *Q7* and *Q9* genes are similar as they are for the *Q6*, *Q8* and *Q10* genes. Thus, alternating gene regions are more similar to one another than to adjacent gene regions, suggesting that after the initial gene duplication, subsequent duplications involved gene pairs. It is not clear whether the duplicated pairs are *Q4* + *Q5*, *Q6* + *Q7*, *Q8* + *Q9* and *Q10* + ? or ? + *Q4*, *Q5* + *Q6*, *Q7* + *Q8* and *Q9* +



Both conserved and polymorphic segments of DNA map to the *Qa2,3* region of the B10 and BALB/c mouse. The restriction maps of the intergenic DNA downstream from *Q5* up to and including *Q10* of the B10 mouse are very similar to the corresponding sequences of cluster 1 from the BALB/c mouse³. Gene *Q10*, which has been sequenced in our laboratory¹¹ and may encode a non-polymorphic, liver-specific¹², secreted¹³ class I antigen, corresponds to the gene in cluster 9 of Steinmetz *et al.*³. Rogers *et al.*¹⁴ have shown independently that the BALB/c clusters 1 and 9 are linked. Comparison reveals that the equivalent of genes *Q8* and *Q9* of the B10 mouse may have undergone an unequal crossing-over in the BALB/c mouse (Fig. 4b) to produce a fusion gene in a manner similar to the Hb-Lepore gene from the human δ - and β -globin genes¹⁵⁻¹⁷. Thus, the restriction map of the DNA flanking gene 7 of cluster 1 of the BALB/c mouse is almost identical to the 5' map of

gene *Q8* and the 3' map of gene *Q9* from the B10 MHC (Fig. 4b). The map of genes *Q4*, *Q6* and *Q7* is almost identical to the corresponding region of the BALB/c MHC (Fig. 4b). It follows, therefore, that gene *Q7* of the B10 mouse is the allele of the 27.1 'pseudogene' (ref. 18) in the BALB/c mouse. The maps of the *Q2*, *Q3* and *Q5* genes differ significantly between the two haplotypes. We have not detected the equivalent of cluster 6 of BALB/c *Qa2,3* region^{3,4} that contains a gene encoding the *Qa2,3* antigen¹⁹.

The B10 *TL* region comprising genes *T1* to *T10* closely parallels clusters 3 and 4 of the BALB/c mouse (Fig. 4c). This suggests that these clusters in fact overlap; alternatively, a deletion may have occurred between genes *T4* and *T5* in the B10 mouse. Genes *T11*–*T13* do not seem to be present in any of the *TL*-mapped clusters of the BALB/c mouse, and we have not detected analogues of the BALB/c clusters 5, 7, 8, 10 and 12 in the B10 genome. The basis for the differential expression of the *TL* antigen in the two haplotypes may well be in these non-homologous regions.

Translocations during evolution

During these studies we have noted several interrelationships between genes located at distant sites that may indicate how the MHC class I gene family has evolved. The 5' flanking *TL* region probe hybridizes to the 5' flanking regions of genes *Q1* to *Q3* in the *Qa2,3* region genes as well as *T5*, *T7* and *T13* of the *TL* region. The similarity of these genes may be due to a gene conversion-like process or a translocation of genes *Q1* to *Q3* from the *TL* region.

A more intriguing relationship is the striking similarity of the

two genes in the *H-2K* region to the *Q6* and *Q7* as well as the *Q8* and *Q9* gene pairs. The genes in each pair are in a head-to-tail configuration and are about the same distance apart. The *H-2K* and *Qa2,3* region gene pairs show similar hybridization patterns with the three 5' flanking *K*-region probes and the 3' flanking *K*-region probe. For example, the 0.64-kb *Bam*HI 5' flanking *K*-region probe hybridizes to 0.64-kb *Bam*HI fragments flanking the *H-2K^b*, *Q6* and *Q8* genes, while it hybridizes to 0.68-kb fragments flanking the other genes in each pair (that is, *K1*, *Q7* and *Q9*). The 3' flanking *K*-region probe also hybridizes to both the *K* region and *Qa2,3* region gene pairs. Together, these data suggest that the *H-2K^b* gene may be related to the *Q6* and *Q8* genes and that the gene *K1* is related to the genes *Q7* and *Q9* and perhaps the *Q5* gene. This similarity probably reflects an evolutionary relationship between the genes. In some animals, all class I gene loci are telomeric to the class II genes. In the mouse, however, the *H-2K* region is centromeric to the class II genes. For this reason it has been suggested that the *H-2K* locus arose by translocation to its present site²⁰. We propose that the *H-2K* region was generated by the translocation of a *Q6*- and *Q7*-like gene pair from the *Qa2,3* region.

L.G. was the recipient of a fellowship from the Helen Hay Whitney Foundation. E.H.W. was initially supported by a fellowship from the Deutsche Forschungsgemeinschaft. We thank Georg Wiedera for preparing the LS library, Neil Malone for preparing cosmid DNA and Madlyn Nathanson for the preparation of this manuscript. We also thank M. Steinmetz for providing *H-2* cDNA probes and B. Jordan for the human genomic probes. Part of this work was supported by the British MRC and part by Biogen NV.

Received 21 February; accepted 25 May 1984.

1. Klein, J. *Biology of the Mouse Histocompatibility Complex* (Springer, Berlin, 1975).
2. Coligan, J. E., Kindt, T. J., Uehara, H., Martinko, J. & Nathenson, S. G. *Nature* **291**, 35–39 (1981).
3. Steinmetz, M., Winoto, A., Minard, K. & Hood, L. *Cell* **28**, 489–498 (1982).
4. Winoto, A., Steinmetz, M. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3425–3429 (1983).
5. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
6. Weiss, E. *et al. EMBO J.* **2**, 453–462 (1983).
7. Mellor, A. L. *et al. Nature* **298**, 529–534 (1982).
8. Oudshoorn-Snoek, M., Demant, P., Mellor, A. L. & Flavell, R. A. *Proc. natn. Acad. Sci. U.S.A.* **15**, 2027–2032 (1983).
9. Townsend, A. R. M., Taylor, P. M., Mellor, A. L. & Askonas, B. A. *Immunogenetics* **17**, 283–294 (1983).
10. Kvist, S., Roberts, L. & Dobberstein, B. *EMBO J.* **2**, 245–254 (1983).
11. Mellor, A. L., Weiss, E. H., Kress, M., Jay, G. & Flavell, R. A. *Cell* **36**, 139–144 (1984).
12. Cosman, D., Kress, M., Khoury, G. & Jay, G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4947–4951 (1982).

13. Cosman, D., Khoury, G. & Jay, G. *Nature* **295**, 73–76 (1982).
14. Rogers, J. H. & Willison, K. R. *Nature* **304**, 549–552 (1983).
15. Baglioni, C. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1880–1886 (1962).
16. Flavell, R. A., Kooter, J. M., de Boer, E., Little, P. F. R. & Williamson, R. *Cell* **15**, 25–41 (1978).
17. Mears, J. G., Ramirez, F., Leibowitz, D. & Bank, A. *Cell* **15**, 15–23 (1978).
18. Steinmetz, M. *et al. Cell* **25**, 683–692 (1981).
19. Goodenow, R. A. *et al. Nature* **300**, 231–237 (1982).
20. Bodmer, W. F. *Tissue Antigens* **17**, 9–20 (1981).
21. Malissen, M., Malissen, B. & Jordan, B. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 893–897 (1982).
22. Steinmetz, M. *et al. Cell* **24**, 125–134 (1981).
23. Grosveld, F. G., Dahl, H. H. M., de Boer, E. & Flavell, R. A. *Gene* **13**, 227–237 (1981).
24. Ish-Horowitz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989–2998 (1981).
25. Grosveld, F. G. *et al. Nucleic Acids Res.* **10**, 6715–6732 (1982).
26. Weiss, E. H. *et al. Nature* **301**, 671–674 (1983).
27. Bernards, R. & Flavell, R. A. *Nucleic Acids Res.* **8**, 1521–1533 (1980).
28. Weislander, L. *Analyt. Biochem.* **98**, 305 (1979).

Functional role for *c-myc* in mitogenic response to platelet-derived growth factor

H. A. Armelin*, M. C. S. Armelin*, K. Kelly†, T. Stewart†, P. Leder†, B. H. Cochran† & C. D. Stiles‡

* Department of Biochemistry, University of Sao Paulo, CP 20780, Sao Paulo, Brazil

† Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Department of Microbiology and Molecular Genetics, Harvard Medical School and the Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA

In BALB/c-3T3 cells, expression of the c-myc gene is stimulated by platelet-derived growth factor (PDGF). Using mouse mammary tumour virus promoter: c-myc recombinant plasmids, 3T3 sublines were constructed in which hydrocortisone was the primary determinant of myc mRNA content. The c-myc gene product is an intracellular mediator of the growth response to PDGF though probably not the only one. Both the human and the mouse c-myc genes stimulate clonal growth of 3T3 cells in PDGF-free medium suggesting new strategies for analysis of oncogenes which do not function in focus formation assays.

PLATELET-derived growth factor (PDGF)^{1–4} renders BALB/c-3T3 cells 'competent' to replicate their DNA and divide⁵. The PDGF-induced competent state consists of an acquired sensitivity to the mitogenic action of epidermal growth factor (EGF) and the insulin-like growth factors^{6,7}. The induction of competence is a transcription-dependent event⁸. At least five

independent members of a PDGF-inducible gene family which we term 'competence' have been isolated by molecular cloning techniques⁹; moreover, *c-myc* seems to be a member of the competence gene family in 3T3 cells¹⁰. The abundance level of *c-myc* mRNA increases more than 40-fold several hours after the addition of PDGF to the culture medium of density-arrested

3T3 cell monolayers. Growth factors such as EGF and insulin have little effect on *c-myc* expression¹⁰.

The observations on induction of *c-myc* by PDGF raise an apparent paradox. The putative structural gene for PDGF is *c-sis* (refs 11,12). Exposure of 3T3 cells to PDGF or transfection with *c-sis* cDNA¹³ results in a strong growth response, together with morphological alterations characteristic of transformation^{14,15}. By contrast, transfection with *c-myc* DNA has never been shown to influence the growth or morphology of 3T3 cells. Conventional assays for transforming activity are conducted with serum-supplemented medium and are contingent upon induction of morphological alterations in an indicator cell line—usually NIH 3T3 (refs 16–18). Loss of specific growth factor requirements may not be tightly coupled to changes in morphology. For this reason we set out to determine whether gratuitously induced expression of *c-myc* within 3T3 cells could override the growth requirement for PDGF.

Construction of BALB/c-3T3 sublines

Recombinant plasmids containing the two coding exons of mouse *c-myc* under control of the mouse mammary tumour virus (MMTV) promoter were constructed. These recombinant plasmids were stably integrated into BALB/c-3T3 cells by co-transfection with a plasmid (HSV-neo) which conferred resistance to the drug G418. Essential features of the recombinant plasmids and the transfection protocol are depicted in Fig. 1. Multiple G418-resistant 3T3 cell colonies from each transfection were isolated and thereafter maintained in G418-supplemented medium. Integration of the exogenous *myc* plasmid constructs into the G418-resistant 3T3 sublines was confirmed by Southern analysis of *Bam*HI- and *Eco*RI-digested cellular DNA.

Expression of the exogenous *myc* plasmids was quantified by *S*₁ nuclease analysis¹⁹ (Fig. 2). As a frame of reference, the abundance of messenger RNA transcripts directed by the exogenous *myc* plasmids was compared with the abundance of *c-myc* transcripts in the parental 3T3 cells incubated in the absence or presence of PDGF. The abundance level of *myc* mRNA transcripts directed by the exogenous plasmids varied widely among different G418-resistant 3T3 sublines; however, we were able to identify two G418-resistant lines (7E and 9E) in which the exogenous *c-myc* gene was expressed at high levels and was steroid responsive (Fig. 2). Scanning densitometry of *S*₁ nuclease assays conducted with the 7E and 9E sublines provides an estimate of *myc* mRNA abundance relative to the parental 3T3 line. As summarized in Table 1, the abundance level of *myc* mRNA transcripts within hydrocortisone-treated 7E and 9E cell cultures exceeds that of the PDGF-treated parental 3T3 cells. Even without hydrocortisone treatment there is substantial constitutive transcription from the exogenous *myc* plasmids; this probably reflects enhancer activity of the strong Harvey sarcoma virus (HSV) promoter carried by the drug resistance marker plasmid together with exogenous activity of the MMTV promoter.

Hydrocortisone stimulation of DNA synthesis

Figure 3 shows the effect of hydrocortisone on PDGF-deprived monolayers of 7E cells. As a control, the effect of hydrocortisone on a control subline (4A) which was transfected with the HSV-neo marker plasmid only is depicted. The nuclear labelling index of the 4A cells is less than 0.5% in the presence or absence of hydrocortisone. Approximately 4% of the 7E cells were labelled with ³H-thymidine in the absence of hydrocortisone. Addition of hydrocortisone increased the nuclear labelling index to ~28% in 7E cell cultures.

As shown in Fig. 4, the differential between the 4A and 7E sublines does not reflect random clonal variation. A number of independent G418-resistant 3T3 cell sublines were isolated and screened for replicative DNA synthesis in the absence of PDGF. Control sublines were transfected with HSV-neo alone or co-transfected with HSV-neo and MMTV-XhoI-*myc* (from which the majority of the *myc* translatable sequences are deleted; see

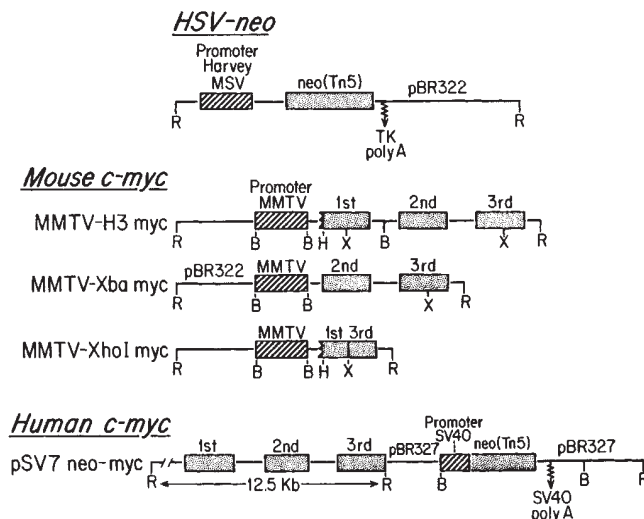


Fig. 1 Recombinant plasmids and transfection protocol. Salient features of the recombinant plasmids are as follows: HSV-neo (from D. Kaplan and T. Roberts, Dana Farber Cancer Institute) contains the Tn5 transposon²¹ activated by the Harvey murine sarcoma virus promoter in pBR322; MMTV-H3 fuses the MMTV promoter to mouse *c-myc* at a *Hind*III site located within the first (nontranslatable) *myc* exon. All sequence 5' from the *Hind*III site, including the normal *myc* promoter elements and TATA box, are thus deleted and transcription initiates from within the MMTV promoter; MMTV-Xba-*myc* is identical to MMTV-H3 except that the entire first exon is deleted; MMTV-XhoI is a non-functional *myc* control in which the translatable sequence from all of the second and most of the third exons is deleted; pSV7-neo-*myc* contains the entire 12.5-kb *Eco*RI fragment of human *c-myc*²² inserted into the pSV7-neo plasmid²¹. Restriction sites and other features are indicated by R (*Eco*RI); B (*Bam*); H (*Hind*III); X (*Xho*I); TK poly A (herpes simplex virus thymidine kinase polyadenylation signal); SV40 poly A (SV40 polyadenylation signal). **Methods:** The MMTV-*myc* plasmids were mixed with HSV-neo at ratios between 40:1 and 25:1 and stably integrated into BALB/c-3T3 by co-transfection followed by selection for resistance to G418²¹. Transfections were carried out in 60-mm dishes by the calcium phosphate precipitation method²³ using 15 µg sheared salmon sperm DNA as carrier, 0.1–0.5 µg HSV-neo plasmid and 2.5–20 µg MMTV-*myc* plasmid. After exposure to the DNA:calcium phosphate coprecipitates, the 3T3 cells were incubated overnight in DMEM supplemented with 10% calf serum and 5 mM of sodium butyrate²⁴. Fresh serum-supplemented DMEM was then added. The cultures were incubated for an additional 48 h and then subcultured at a 1:5 or 1:10 ratio into DMEM supplemented with 10% calf serum and 0.5 mg ml⁻¹ G418 (Gibco). The cultures were incubated for 2–3 weeks with a medium change at 2–3-day intervals. At this time, well isolated G418-resistant colonies were picked and subcultured. Transfections with pSV7-neo-*myc* did not require the HSV-neo marker plasmid but were otherwise identical. Control sublines were transfected with the HSV-neo marker plasmid alone or with pSV7-neo alone.

Fig. 1). None of the nine control sublines showed appreciable nuclear labelling in the absence of PDGF and none showed a response to hydrocortisone (Fig. 4). Five out of 9 sublines co-transfected with MMTV-Xba-*myc* (including 7E) and 4 out of 10 sublines co-transfected with MMTV-H3-*myc* (including 9E) showed enhanced nuclear labelling in the absence of PDGF; the labelling index of these sublines was further enhanced by the addition of hydrocortisone (Fig. 4).

Response to EGF

As shown in Figs 3 and 4, expression of the *myc* gene increases the fraction of 3T3 cells which incorporate ³H-thymidine in the absence of PDGF. However, the nuclear labelling index of the *myc* transfectants is relatively low. In the absence of EGF and insulin-like growth factors, PDGF is an inefficient mitogen for 3T3 cells. Previous work has shown that PDGF interacts synergistically with EGF and the insulin-like growth factors to give

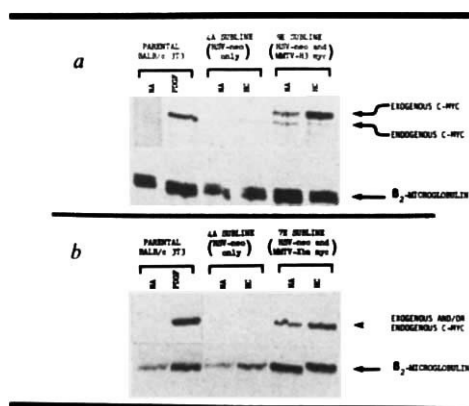


Fig. 2 Induction of *c-myc* by hydrocortisone in 3T3 cell sublines transfected with MMTV-*myc* recombinant plasmids. The parental BALB/c-3T3 cells and three G418-resistant sublines, transfected as indicated, were grown to the confluent monolayer stage in 150-mm culture dishes using DMEM and 10% calf serum. At this point, the serum-supplemented medium was removed and replaced with DMEM supplemented with 5% platelet-poor plasma. After 24 h in plasma-supplemented medium (which is deficient in PDGF), the parental 3T3 cells were exposed for 3 h to a saturating dose of partially purified PDGF¹ (400 U ml⁻¹, equivalent to 80 ng ml⁻¹) to display the induction of the endogenous *c-myc*. The G418-resistant sublines were exposed for 3 h to 100 ng ml⁻¹ hydrocortisone. Control cultures of the parental cells and of each subline were left untreated (NA). Total cellular RNA was collected as described previously¹⁰ and the abundance level of *myc* transcripts was displayed by S₁ nuclease analysis. In panel *a*, the probe used was a fragment of the *myc* first exon described previously¹⁰. With this probe, transcripts originating from the (endogenous) 3T3 cell *myc* and the (exogenous) MMTV-H3-*myc* can be distinguished from each other. The latter transcripts originate from within the MMTV promoter and are 24 bases longer than the predominant form of transcript originating from the endogenous *myc* promoter¹⁰. In panel *b*, the probe used was a fragment of the *myc* second exon. This probe cannot distinguish between transcripts of endogenous and exogenous origin; however, it was necessary for analysis of the 7E subline and other sublines transfected with MMTV-Xba-*myc* in which the entire first exon is deleted (Fig. 1). A parallel S₁ nuclease assay was conducted on each sample using a probe for β_2 -microglobulin as an internal standard for RNA content¹⁰.

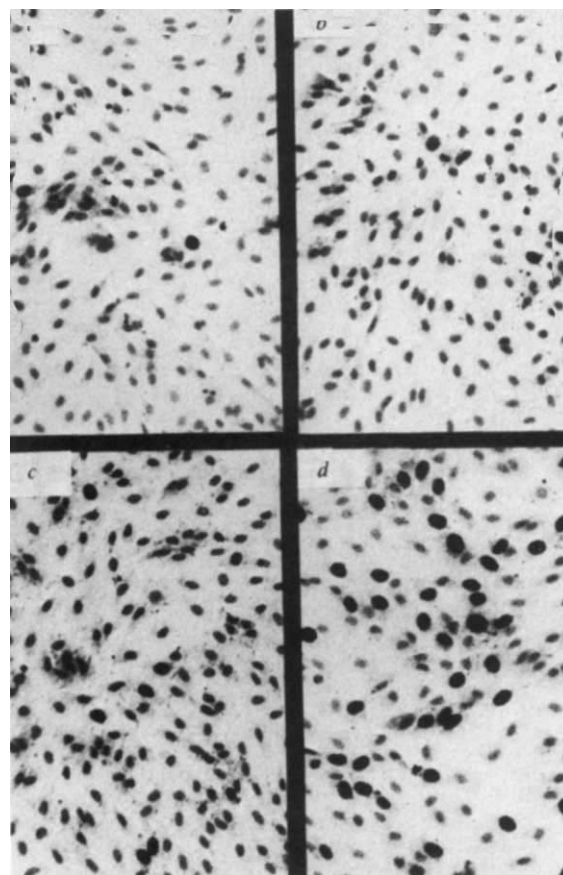


Fig. 3 Stimulation of DNA synthesis by hydrocortisone in 3T3 sublines transfected with an MMTV-*myc* recombinant plasmid. Density-arrested, PDGF-deprived cultures of the 4A subline (*a*, *b*) and the 7E subline (*c*, *d*) were obtained as described in Fig. 2 legend. Hydrocortisone (100 ng ml⁻¹) and ³H-thymidine (2 μ Ci ml⁻¹; 5×10^{-7} M) were added. After 24 h, the cells were fixed, processed for autoradiography and counterstained with Giemsa. The per cent of cells with labelled nuclei in the 4A cultures without (*a*) or with (*b*) hydrocortisone is less than 0.5%. In the absence of hydrocortisone, 4% of the 7E cells show ³H-thymidine uptake (*c*). In the presence of hydrocortisone (*d*), 28% of the 7E cultures were labelled in this experiment.

Table 1 Relative abundance of *c-myc* mRNA

	Parental BALB/c-3T3	4A subline HSV-neo only	7E subline HSV-neo and MMTV-Xba- <i>myc</i>	9E subline HSV-neo and MMTV-H3- <i>myc</i>
NA	1	1.5	13	39
PDGF	40	ND	ND	ND
Hydrocortisone	ND	2	90	175

The S₁ nuclease assays depicted in Fig. 2 were analysed by scanning densitometry. Results are summarized as *myc* abundance units relative to parental BALB/c-3T3 cells in PDGF-free medium. NA, untreated control cultures; ND, not determined.

an optimum growth response^{6,7}. We reasoned, therefore, that any growth-promoting effects of *myc* expression could be amplified by the addition of EGF and/or insulin to the culture medium.

As shown in Fig. 5, density-arrested 9E cells responded to EGF by initiating DNA synthesis in serum-free medium supplemented only with insulin. Hydrocortisone enhanced the effect of EGF, presumably because it further increases *myc* expression (Fig. 2). On the other hand, the control 4A subline in the same conditions showed no response to EGF or EGF plus hydrocortisone, but instead required the combination of PDGF and EGF for an efficient mitogenic response (Fig. 5). These responses to EGF and hydrocortisone are reflected by an increase in cell number during long-term experiments (Table 2). Again, these effects do not reflect random clonal variations. None of the 9

control sublines showed this response to EGF and hydrocortisone whereas 5 of 9 MMTV-H3 sublines and 4 of 10 MMTV-Xba sublines displayed the response (data not shown). Neither the endogenous nor the exogenous *myc* gene was induced by EGF treatment in 9E cell cultures as shown by S₁ nuclease assay (data not shown).

Direct selection of *myc* transfected 3T3 cells

When expression of the *myc* gene is placed under control of viral promoters, the growth requirement for PDGF is diminished (Figs 3–5, Table 2). The selection protocol shown schematically in Fig. 6 illustrates a practical consequence of this. BALB/c-3T3 cells were co-transfected with a neomycin-resistance marker and any of several *myc* gene constructs. After 48–72 h incubation in serum-supplemented medium, the transfected cells were selected

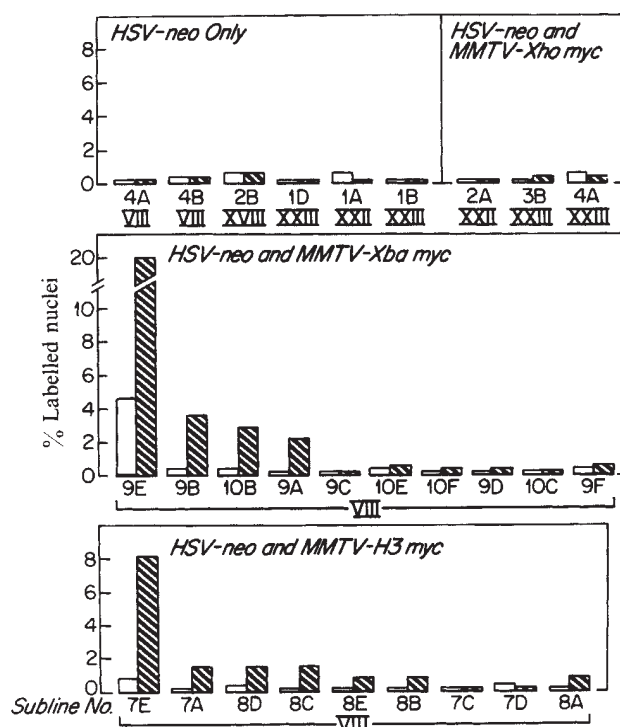


Fig. 4 Stimulation of replicative DNA synthesis by mouse *c-myc* in randomly isolated, G418-resistant 3T3 sublines. A functional mouse *c-myc* plasmid construct (MMTV-H3-myc or MMTV-Xba-myc) was transfected into 3T3 sublines together with the HSV-neo marker plasmid as described in Fig. 1 legend. Control cultures were transfected with HSV-neo alone or HSV-neo and MMTV-XhoI-myc. Density-arrested, PDGF-deprived cultures were prepared as described for Fig. 2 except that microtitre wells were used. The growth-arrested cells, in plasma-supplemented DMEM, were incubated overnight with ^3H -thymidine ($2\mu\text{Ci ml}^{-1}$; $5 \times 10^{-7} \text{ M}$) and 100 ng ml^{-1} hydrocortisone or a solvent control. After 24 h, the cultures were fixed, processed for autoradiography and scored for labelled nuclei. Values shown are averages of duplicate cultures.

for growth in culture medium supplemented with G418 and either serum (which contains PDGF) or platelet-poor plasma (PDGF-deficient). The ratio of colonies formed in plasma-supplemented medium to colonies formed in serum-supplemented medium reflects expression of a functional *myc* gene product. As experimental controls, cells were transfected with the neomycin-resistance marker plasmid alone or co-transfected with the marker plasmid and MMTV-XhoI-myc.

Table 3a summarizes the results with various mouse *myc* constructs. In control transfections with HSV-neo only, the plasma:serum ratio was low (0.13). Identical results (plasma:serum ratio 0.10) were obtained with co-transfection of HSV-neo and MMTV-XhoI-myc. Co-transfection with HSV-

Table 2 Effects of EGF and hydrocortisone on growth of 4A and 9E sublines in PDGF-free medium

Additions to culture medium	4A subline	9E subline
None	3.6	4.9
Hydrocortisone	3.7	7.5
EGF	7.1	24.0
Hydrocortisone + EGF	6.1	25.0
PDGF	10.6	15.0

Cells from the control 4A subline or the MMTV-H3-myc 9E subline (see Fig. 2) were plated in 10% calf serum Dulbecco's modified Eagle's medium (DMEM) at 1.2×10^4 cells cm^2 in 35-mm dishes. After 24 h the medium was removed. The cultures were washed with phosphate-buffered saline and incubated in fresh 10% plasma and DMEM for another 24 h. Additions were then made of hydrocortisone (100 ng ml^{-1}), EGF (1 ng ml^{-1}) or PDGF (crude preparation 25 U ml^{-1}) as indicated. Medium was changed after 2 days. Cells were collected and counted on day 4 following additions. Values shown are cells per cm^2 ($\times 10^{-4}$). Each value is the average of duplicate plates; deviations no greater than 5%.

neo and either MMTV-H3-myc or MMTV-Xba-myc increased the plasma:serum ratio by a factor of five (0.57 and 0.52 respectively).

Table 3b summarizes transfections using a single plasmid (pSV7-neo-myc) which contained both the neomycin-resistance marker and the complete human *myc* gene (pSV7-neo-myc; see Fig. 1). The results are similar to those obtained with mouse *myc*. Transfections with the human *myc* gene increased the plasma:serum ratio 6.3-fold relative to transfection with the identical plasmid (pSV7-neo) without *myc* (plasma:serum ratios 0.44 and 0.07 respectively). For comparison, Table 3b also presents data from transfection with a neo-v-sis plasmid. As expected, v-sis was very effective and raised the plasma:serum ratio by a factor of 8 above the background level.

The ability to develop colonies in plasma was also analysed for randomly isolated stable transfectant sublines. The results are summarized in Fig. 7. The G418-resistant sublines fell into three distinct classes: v-sis transfectants display the same plating efficiency in plasma and serum whereas the controls (HSV-neo transfectants) grow poorly in plasma. The *myc* transfectants exhibited intermediate growth in plasma. The data are consistent with the view that *myc* expression enhances colony development in plasma; however, the effect of *myc* is weaker than the effect of *sis* (see below).

Discussion

The product of the *c-sis* gene, PDGF^{11,12}, stimulates expression of the *c-myc* gene in 3T3 cell cultures¹⁰. Our aim here is to determine whether the *c-myc* gene product functions as an intracellular mediator of the mitogenic response to PDGF. When the abundance level of *myc* mRNA is gratuitously induced in 3T3 cells by the MMTV promoter, both the probability of cell division in the absence of PDGF and the sensitivity to the

Table 3 PDGF-free selection assays with HSV-neo and MMTV-myc recombinant plasmids (a) and with human *c-myc* and monkey v-sis (b)

Plasmids transfected	No. of colonies in serum	No. of colonies in plasma	Ratio plasma:serum	Total no. of experiments
a HSV-neo only	209	28	0.13	16
HSV-neo and MMTV-Xho-myc	49	5	0.10	7
HSV-neo and MMTV-Xba-myc	108	62	0.57	6
HSV-neo and MMTV-H3-myc	165	85	0.52	12
b pSV7-neo	904	62	0.07	3
pSV7-neo-myc	1,602	702	0.44	4
pSV7-neo-v-sis	314	176	0.56	4

a, BALB/c-3T3 cells were transfected with HSV-neo and with various MMTV-myc recombinant plasmids as indicated. The transfected cells were then carried through the selection protocol outlined in Fig. 6. b, BALB/c-3T3 cells were transfected with human *c-myc* (pSV7-neo-myc; see Fig. 1) or with the corresponding parental marker plasmid (pSV7-neo). Additionally, some transfections were conducted with the monkey v-sis gene carried in the pSV2-neo plasmid which is similar to pSV7-neo. The pSV2-neo-v-sis plasmid was derived from the C60 clone of simian sarcoma virus described by Gelman *et al.*²⁰.

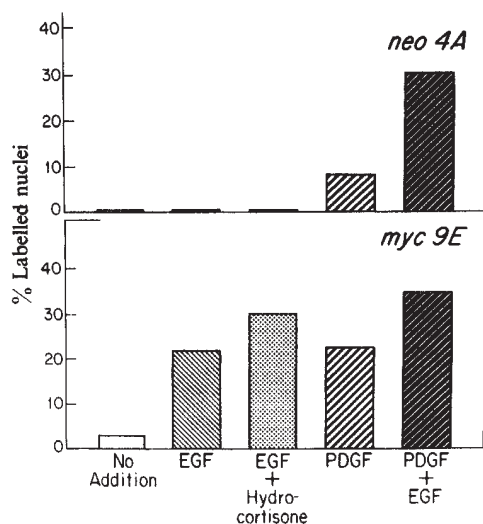


Fig. 5 Sensitization to EGF by constitutive expression of *c-myc*. Density-arrested, PDGF-deprived cultures of the 9E subline (see Fig. 2) were prepared in microtitre wells as described in Fig. 2 legend. The cell monolayers were washed with phosphate-buffered saline. Fresh DMEM supplemented with insulin ($5 \mu\text{g ml}^{-1}$), EGF (10 ng ml^{-1}), hydrocortisone (100 ng ml^{-1}) or partially purified PDGF (20 U ml^{-1}) were added singly or in combinations as indicated. The cultures were incubated for 24 h in the presence of ^3H -thymidine ($2 \mu\text{Ci ml}^{-1}$; 10^{-7} M) and then fixed and processed for autoradiography. 500–1,000 cells were counted per well. The values indicated are averages of duplicate cultures.

mitogenic activity of EGF are increased. However, this growth response is not directly proportional to the intracellular content of *myc* mRNA. The 7E and 9E sublines can be induced to express more *myc* mRNA in the absence of PDGF than the parental 3T3 cells express in the presence of PDGF (Fig. 2, Table 1); yet only a fraction of the 7E and 9E cell populations replicate DNA in a 24-h period as opposed to almost 100% of the PDGF-stimulated parental cells in similar conditions. In addition, the characteristic 'cobblestone street' morphology of a confluent 3T3 monolayer is unaltered in hydrocortisone-treated 7E and 9E sublines (see Fig. 3). By contrast, PDGF-treated 3T3 cells, containing similar levels of *myc* mRNA, elongate and become fusiform^{14,15}. Similarly, 3T3 cells transfected with *c-sis* cDNA exhibit morphological alterations¹³. In summary, the abundance level of *myc* mRNA within 3T3 cells can be increased either by exposure to PDGF or by viral promoter; however, the consequences for 3T3 cell growth and morphology are more radical when *myc* is expressed in response to PDGF.

There are several potential explanations for the differential response of 3T3 cells to PDGF-induced *myc* and gratuitously induced *myc*. It is conceivable, for example, that expression of *c-myc* protein does not correlate with observed mRNA levels. A more interesting possibility is that the *myc* gene product functions coordinately with the other PDGF-induced gene products to mediate the growth response within the cell. We have recently found a striking homology between the nucleotide sequence of a PDGF-inducible gene isolated previously by cloning techniques⁹ and the *c-fos* oncogene (B.H.C., J. Zullo and C.D.S., manuscript in preparation). Some or all of the other PDGF-inducible competence genes may be hitherto undiscovered oncogenes which contribute to the PDGF-induced mitogenic response.

The growth-promoting action of *myc* within 3T3 cells can be conveniently displayed in a double selection protocol which is contingent upon cell colony formation in the absence of PDGF. The double selection procedure may be of value in structure/function analysis of oncogene products. The procedure saves time by providing a direct assay for oncogene biological activity; moreover, the assay is objective. Only the number of

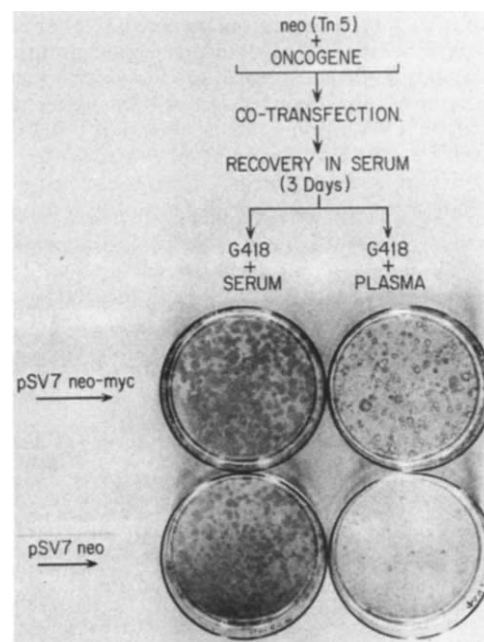


Fig. 6 Growth-promoting action of *c-myc* displayed in a double selection protocol. BALB/c 3T3 cells (clone A31) were exposed to calcium phosphate:DNA co-precipitates containing the HSV-neo plasmid and various mouse *c-myc* constructs as indicated in Fig. 1. Alternatively, the human *c-myc* gene (in the pSV7-neo plasmid) and the monkey *v-sis* gene (in the pSV2-neo plasmid) were introduced into 3T3 cells without the HSV-neo plasmid. Following exposure to the calcium phosphate:DNA co-precipitates, the 3T3 cells were allowed to proliferate for 3 days in DMEM supplemented with 10% calf serum. At this time, the cells were collected by trypsin digestion and subcultured into 100-mm diameter culture dishes which had been coated with fibronectin (Collaborative Research). Fibronectin-coated culture dishes were essential for good results because transfected cells tend to become detached from the plates on prolonged culture in plasma-supplemented medium. The cultures were then incubated in DMEM supplemented with G418 (Gibco) and either 10% calf serum or 10% human platelet-poor plasma. After 2–3 weeks incubation, with medium change at 2–3-day intervals, the cultures were fixed in phosphate-buffered 10% formalin and stained with crystal violet. The inset shows a set of cultures which were transfected with the human *c-myc* gene (pSV7-neo-myc) or with the control marker plasmid (pSV7-neo) alone.

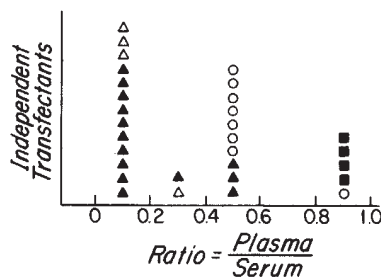


Fig. 7 Relative plating efficiency (plasma:serum) of randomly isolated 3T3 sublines transfected with *myc*, *sis* or HSV-neo. ▲, Neo only (HSV-neo); △, neo + deleted *c-myc*; ○, neo + mouse *c-myc*; ■, neo + *v-sis*. Stock cultures of randomly isolated G418-resistant 3T3 sublines were collected by trypsin digestion, counted in a haemocytometer and diluted. The cells were plated at a density of 500 per dish into 60-mm culture dishes which had been precoated with fibronectin. The cultures were incubated for 10–14 days in medium supplemented with G418 and either 10% serum or 10% plasma. The cultures were then fixed with 10% formalin in phosphate-buffered saline, stained with crystal violet and counted.

colonies to grow in the absence of a specific growth factor need be determined; morphological characteristics of the colonies are not a factor. Finally, the double selection protocol, using culture medium depleted of other specific growth factors, has the potential to display the effects of cell oncogenes which are not revealed by conventional 'focus formation' assays.

We thank Drs Jim Battey and Hunt Potter for their gift of the pSV7-neo-myc clone. Megan Callahan constructed the MMTV-Xho I myc clone and provided technical assistance. This

work was supported by grants from the American Business for Cancer Foundation, the American Cancer Society (Research Development Program), the E.I. DuPont de Nemours and Co., and the NIH. H.A.A. was a Fellow of the Guggenheim Foundation. M.C.S.A. was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil. B.H.C. was supported by a predoctoral fellowship from the Noble Foundation. C.D.S. is supported by a Faculty Research Award from the American Cancer Society.

Received 26 June; accepted 6 July 1984.

1. Antoniadis, H. N., Scher, C. D. & Stiles, C. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1809–1813 (1979).
2. Heldin, C. H., Westermark, B. & Wasteson, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3722–3726 (1979).
3. Deuel, T. et al. *J. biol. Chem.* **256**, 8896–8899 (1981).
4. Raines, E. W. & Ross, R. J. *biol. Chem.* **257**, 5154–5160 (1982).
5. Pledger, W. J., Stiles, C. D., Antoniadis, H. N. & Scher, C. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4481–4485 (1977).
6. Stiles, C. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1279–1283 (1979).
7. Leof, E. B., Wharton, W., Van Wyk, J. J. & Pledger, W. J. *Expl Cell Res.* **141**, 107–115 (1982).
8. Smith, J. C. & Stiles, C. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4363–4367 (1981).
9. Cochran, B. H., Reffel, A. C. & Stiles, C. D. *Cell* **33**, 939–947 (1983).
10. Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 603–610 (1983).

11. Doolittle, R. F. et al. *Science* **221**, 275–277 (1983).
12. Waterfield, M. D. et al. *Nature* **304**, 35–39 (1983).
13. Clarke, M. F. et al. *Nature* **308**, 464–467 (1984).
14. Mellstrom, K. et al. *J. Muscle Res. Cell Motility* **4**, 589–609 (1983).
15. Bockus, B. J. & Stiles, C. D. *Expl Cell Res.* **153**, 186–197 (1984).
16. Shih, C., Pahdy, L. C., Murray, M. & Weinberg, R. A. *Nature* **290**, 261–264 (1981).
17. Krontis, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1181–1184 (1981).
18. Peruchio, M. et al. *Cell* **27**, 467–476 (1981).
19. Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
20. Gelmann, E. P., Wong-Staal, F., Kramer, A. R. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3373–3377 (1981).
21. Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327–341 (1982).
22. Battey, J. et al. *Cell* **34**, 779–789 (1983).
23. Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456–467 (1973).
24. Gorman, C. M., Howard, B. H. & Reeves, R. *Nucleic Acids Res.* **11**, 7631–7648 (1983).

LETTERS TO NATURE

Infrared spectropolarimetry of Seyfert galaxy NGC1068

D. K. Aitken*, J. A. Bailey†, G. Briggs*, J. H. Hough† & P. F. Roche†

* Department of Physics (RAAF Academy), University of Melbourne, Parkville 3052, Victoria, Australia

† Anglo-Australian Observatory, PO Box 296 Epping, NSW 2121, Australia

‡ Physics Department, Hatfield Polytechnic, Hatfield, Hertfordshire AL10 9AB, UK

Although most of the luminosity of Seyfert galaxy nuclei is emitted in the IR, the relative amounts of thermal dust and non-thermal emission are uncertain. Seyfert galaxies generally have smooth and featureless IR spectra, in contrast to the prominent dust emission features seen in galaxies whose nuclei resemble giant H II regions—the star burst nuclei^{1,2}. Nevertheless, in type 2 Seyferts, the general form of the energy distribution and lack of marked variability suggest that here too thermal dust emission is dominant^{3–5}, and in NGC1068 this supposition is supported by the resolved size of its nucleus⁶. Dust is observed in absorption at 10 μm (refs 7, 8) and in emission near 20 μm (ref. 9). Polarization has been detected from the UV to 10 μm , and attributed to scattering and absorption at short wavelengths¹⁰ with a possible non-thermal component^{11,12}, to scattering and non-thermal radiation in the near IR⁹, and to either absorption or non-thermal emission in the 10 μm region¹³. We present here spectropolarimetric measurements of NGC1068, between 8 and 13 μm , which show that the polarized flux at these wavelengths is intrinsic to the emission mechanism and not due to absorption by aligned grains.

The observations were made on 20 August 1983 at the f/36 chopping secondary focus of the 3.9-m Anglo-Australian Telescope (A-AT) and on 13 December on the 3.0 m Infrared Telescope Facility (IRTF) in Hawaii. The UCL (University College, London) grating spectrometer was modified for spectropolarimetry by the insertion of a rotating half-wave plate (Cleveland Crystal Inc., Ohio) on the telescope axis and a fixed wire grid analyser, of efficiency >99%, on the cold surface of the spectrometer close to the field stop. The half-wave plate is not

achromatic and there is a reduction in efficiency at 13 μm , particularly near 8 μm . These effects were calibrated by reference to an external wire grid polarizer and appropriate corrections applied to the data.

Integrations of the 25 detector outputs were recorded at 1 s intervals at each of 16 equispaced positions in a full rotation of the half-wave plate. This sequence of data points was repeated with the source in the other (negative) beam, and a sequence (+ – +) formed a single unit in the usual photometric procedure; the total integration time was 90 min on each telescope. From these data the Stokes parameters for each of the 25 detectors were derived, and to improve the signal-to-noise ratio have been averaged in pairs; the data are presented with a resolution of 0.44 μm . Flux calibration and correction of intensity for telluric absorption were from observations of nearby stars while observation of emission lines in planetary nebulae and H II regions gave the wavelength calibration. For these observations a circular aperture of 4.2- and 5.6-arc s diameter and beam throw 50- and 15-arc s NS were used on the AAT and IRTF respectively; our observations indicate that the flux in the reference beams is <1% of that observed.

A second wire grid was placed on the telescope axis to determine the position angle calibration. This gave the position angle of 10 μm polarization in the BN object in Orion as 118°, a value also found by R. W. Capps (personal communication); we estimate the absolute error on this calibration to be <1°.

The instrumental polarization was found from observation of bright stars. On both telescopes this was apparently constant in magnitude and position angle through the spectrum and of amount $0.25 \pm 0.08\%$; the present data have been corrected for these effects.

The spectral characteristics of polarization and position angle are similar for the two observations of NGC1068, but although the spectrum averaged polarizations agreed very well, the corresponding position angles differed substantially between August and December with respective values $44.1 \pm 3.2^\circ$ and $59.4 \pm 2.2^\circ$. The errors are statistical and may be doubled by uncertainties in the instrumental corrections. Figure 1 shows flux, polarization, position angle and polarized intensity as a function of wavelength after combining the data sets. The spectrum is similar to and consistent with that reported in ref. 8, showing a smooth or power law spectrum subject to silicate-like absorption with an optical depth $\tau_{9.7} \approx 0.50$. This optical depth need not apply to the whole of the 10- μm 1 arc s source resolved by Becklin et

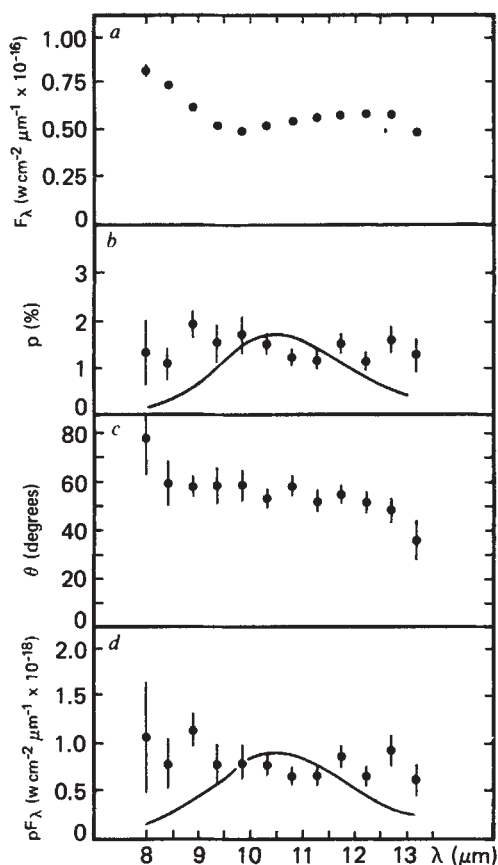


Fig. 1 *a*, Flux, F_{λ} ; *b*, polarization, p ; *c*, position angle, θ ; *d*, polarized intensity, pF_{λ} , of NGC1068 between 8 and 13 μm . The solid curve in *b* is the polarization to the BN object in Orion taken during the same observing period, and in *d* subject to silicate absorption with $\tau_{9.7} = 0.5$.

*al.*⁶, as the spectrum is also consistent with a superposition of smooth spectrum sources with a wide range of optical depths. Also, a small fraction of the emission, up to 20%, could be supplied by optically-thin silicates which, with an effective temperature of 175 K, could account for the observed 20 μm feature. The polarization is roughly uniform through the spectrum with average value $1.39 \pm 0.09\%$. The position angle is also roughly independent of wavelength with a formal average of $54.8 \pm 1.9^\circ$; there may be a slight increase near 8 μm and decrease near 13 μm . The polarization is similar to that observed by Knacke and Capps¹³ in 1974 but our position angles appear to be inconsistent with their measurement of $84^\circ \pm 7^\circ$ at *N*. Knacke and Capps also obtained discrepant polarization angles, and had discounted a previous measurement of 55° made a few months earlier, as due to calibration errors. Lebofsky *et al.*⁹ found significant disagreement with the 3.5- μm position angles reported by Knacke and Capps, and found that the position angle increases with wavelength through *JHKL'* reaching 125° in the latter band; more recent *JHKL'* polarimetry (J.A.B., D. Axon, J.H.H., M. Ward, I. S. McLean and S. R. Heathcote, in preparation) confirms these observations. The observed discrepancies may reflect undetermined systematic errors which are significant at these small polarizations. Nevertheless, there is a distinct possibility that the IR polarization may be variable.

The essential independence of position angle with wavelength between 8 and 13 μm suggests that only a single polarization mechanism is operating in this region, or that any other mechanisms have similar position angles. In other sources, which show 10- μm polarization, the mechanism appears to be differential absorption by aligned non-spherical grains, as for instance, in the BN object¹⁴⁻¹⁶. If the grains are aligned, then there will be

two optical depths, $\tau_{\parallel}$ and $\tau_{\perp}$, appropriate to electric vectors parallel and perpendicular to long and short grain axes. $\tau_{\parallel} - \tau_{\perp} = \Delta\tau$ where $\Delta\tau$ depends on the grain shape and size, complex refractive index, the degree of alignment and its component projected on the sky; usually $\tau_{\parallel} \geq \tau_{\perp}$. In normal Davis-Greenstein mechanism¹⁷, the short grain axes align along the magnetic field and electric vectors parallel to the component of field in the plane of the sky will be preferentially transmitted; many obscured stars show polarization with electric vector roughly aligned with the galactic plane. The polarization (see ref. 18)

$$p = (e^{-\tau_{\perp}} - e^{-\tau_{\parallel}}) / (e^{-\tau_{\perp}} + e^{-\tau_{\parallel}}) \\ = (1 - e^{-\Delta\tau}) / (1 + e^{-\Delta\tau}) = \tanh(\Delta\tau/2) \approx \Delta\tau/2$$

provided $\Delta\tau$ is small; if the band strength of any resonance is weak, $\Delta\tau$ and τ will have a similar wavelength dependence¹⁹. The best studied source which shows substantial IR polarization is the BN object in Orion which shows a broad polarization peak similar to the silicate feature, suggesting a weak band strength and interpreted as due to amorphous silicates¹⁹. Eight more obscured sources showing polarization have been observed and show very similar behaviour (ref. 20 and our work, in preparation). Although the 10 μm spectrum of NGC1068 shows clear evidence for the silicate absorption feature, and the observed position angle of polarization is approximately normal to the position angle of the rotation axis of the galaxy (145°), the wavelength dependence of polarization is unlike that seen in other obscured sources. Figure 1 shows the BN polarization curve superimposed on NGC1068; fitting this curve plus a featureless component to the polarization fraction indicates that $<18\%$ of the polarization can be due to aligned BN-like grains. Absorption in featureless material is unlikely because the polarization should significantly decrease between 8 and 13 μm and a much greater extinction than estimated from the depth of the silicate feature would be necessary.

Lebofsky *et al.*⁹ have considered a model for NGC1068 in which the emission from an intrinsically polarized power law source is scattered and absorbed by surrounding dust. In this way they account for the position angle rotation observed between 1.25 and 3.45 μm . They consider that the nuclear source may be non-thermal and only directly observable between 1 and 5 μm ; scattered light dominates in the visible while at longer wavelengths it is masked by thermal dust emission. Short-term variability would be strong evidence for a non-thermal source and the lack of reproducibility of position angle warrants further study. If the changes apparent in position angles of polarization at 3.5 and 10 μm over periods of the order of months or years are real, then the 10- μm polarization could not be due to aligned grains. The far IR cutoff places severe restrictions on the size of a non-thermal source but Becklin *et al.*⁶ consider that up to 30% of the 10 μm emission could arise from a spatially unresolved object. There is evidence for a bright central object of <0.2 arc s at 2.2 μm (ref. 21), of 0.07 arc s in the radio at 2 cm (ref. 22), and <0.03 arc s in the visible (ref. 23). The latter source is unlikely to suffer the large extinction ($A_v \sim 7$ mag) implied by the 10- μm spectrum, but this extinction need not cover all the sources in the nuclear region. Assuming then that these observations represent a single object which also contributes the 10- μm polarization, its intrinsic polarization must be at least 4.0%, a value similar to the 2.2- μm polarization observed by Lebofsky *et al.*⁹. However a serious difficulty with this interpretation is the large change of position angle between *L'* and *N*, which cannot be explained in terms of Faraday rotation¹³, because it requires twice as large a change between 8 and 13 μm as between *L'* and 8 μm .

Alternatively, we can consider the possibility that the 10- μm polarization is due to thermal emission from aligned grains. Apart from a geometric factor the polarized intensity will be: $pF_{\lambda} \approx \Delta\tau B(T_{\text{grain}})/2$ (see refs 18,19) with wavelength dependence similar to the polarization fraction produced by absorption by aligned grains. Again there is no sign of BN-like structure even when allowance is made for extinction by unaligned silicate

grains of optical depth 0.5. Dilution by featureless emission as implied by a simple two-component model does not affect the polarized intensity, provided it is itself unpolarized. Although it is possible that the 8–13 μm spectrum may be due to radiative transfer effects in a source of moderate silicate optical depth^{9,24}, we consider that the polarization cannot arise in this way because numerical modelling of a range of situations of this sort shows that a strong spectral dependence, or even reversal, is expected.

It seems that if dust emission is responsible for the 10- μm polarization, the grains must be non-silicate and featureless, with temperatures ~ 300 K; their long axes will be predominantly parallel to the electric vector. For these warm grains photon emission can dominate the effects of gas bombardment^{25,26}. Provided the gas density, $n_{\text{gas}} < 10^4 \text{ cm}^{-3}$, and temperature $T_{\text{gas}} < 10^4$ K, the rotation temperature of grains will be less than their internal temperature, and a magnetic field would tend to align grains with their long axes along the field direction, which in this case would lie in the plane of the galaxy. For significant alignment the field strength would have to be $> 10^{-4}$ G. Alternatively, streaming of dust grains with respect to gas along the galaxy plane and possibly driven by the luminous central source could produce the observed position angle^{27,28}. Such grains could also account for the nearly orthogonal position angle at L' , if they obscure a hot inner component, while scattering could give rise to the position angle change in the near IR, as in the model of Lebofsky *et al.* The visual extinction would have to be ≥ 7 mag, additional to the silicate absorption, to produce the observed polarization at K . A prediction from this model is that the polarization at M should be small with position angle close to either 55° or 145° .

These observations indicate that the 10- μm polarization in NGC1068 is not due to absorption by aligned grains but arises in emission. We cannot distinguish between a small and variable contribution from a non-thermal source, or emission from aligned non-silicate grains; position angle monitoring at these wavelengths may be of value. Because other Seyferts are all fainter by more than an order of magnitude, this may well be the only Seyfert which can be studied in this way from ground-based telescopes.

We thank Graeme Holloway and Greg Cherominek of Melbourne University for help before the observations, and the staff of the Anglo-Australian Observatory and the IRTF. This work was supported in part by an ARGS grant. D.K.A., J.A.B., J.H.H. and P.F.R. are visiting astronomers at the IRTF which is operated by the University of Hawaii under contract to NASA.

Received 19 March; accepted 26 June 1984.

- Phillips, M. M., Aitken, D. K. & Roche, P. F. *Mon. Not. R. astr. Soc.* **207**, 25–34 (1984).
- Roche, P. F., Aitken, D. K., Phillips, M. M., & Whitmore, B. *Mon. Not. R. astr. Soc.* **207**, 35–46 (1984).
- Soifer, B. T. & Neugebauer, G. *Infrared Astronomy, IAU Symp. No. 96* 329 (1981).
- Smith, M. G. *16th Symp. Galactic and Extragalactic Infrared Spectroscopy*, 369 (1982).
- Stein, W. A. & Soifer, B. T. *A. Rev. Astr. Astrophys.* **21**, 177 (1983).
- Becklin, E. E., Mathews, K., Neugebauer, G. & Wynn-Williams, C. G. *Astrophys. J. Lett.* **186**, L69–L72 (1973).
- Kleinman, D. E., Gillett, F. C. & Wright, E. L. *Astrophys. J.* **208**, 42 (1976).
- Cutri, R. M. *et al. Astrophys. J.* **245**, 818–828 (1981).
- Lebofsky, M. J., Rieke, G. H. & Kemp, J. C. *Astrophys. J.* **222**, 95–99 (1978).
- Angel, J. R. P., Stockman, H. S., Woolf, N. J., Beaver, A. E. & Martin, P. G. *Astrophys. J. Lett.* **206**, L5–L10 (1976).
- McClean, I. S., Aspin, C., Heathcote, S. R. & McCaughrean, M. J. *Nature* **304**, 609–611 (1983).
- Miller, J. S. & Antonucci, R. R. J. *Astrophys. J. Lett.* **271**, L7–L11 (1983).
- Knacke, R. F. & Capps, R. W. *Astrophys. J. Lett.* **192**, L19–L22 (1974).
- Dyck, H. M., Capps, R. W., Forrest, W. J. & Gillett, F. C. *Astrophys. J. Lett.* **183**, L99–L102 (1973).
- Dyck, H. M. & Beichman, C. A. *Astrophys. J.* **194**, 57–64 (1974).
- Capps, R. W., Gillett, F. C. & Knacke, R. F. *Astrophys. J.* **226**, 863–868 (1978).
- Davis, L. & Greenstein, J. L. *Astrophys. J.* **114**, 206 (1951).
- Dennison, B. *Astrophys. J.* **215**, 529–532 (1977).
- Martin, P. G. *Astrophys. J.* **202**, 393–399 (1975).
- Dyck, H. M. & Lonsdale, C. J. *Infrared Astronomy, IAU Symp. 96* 223–236 (1981).
- McCarthy, D. W., Low, F. J., Kleinman, S. G. & Gillett, F. C. *Astrophys. J. Lett.* **257**, L7–L11 (1982).
- van de Hulst, J. M., Hummel, E. & Dickey, J. M. *Astrophys. J.* **261**, 159–163 (1982).
- Meaburn, J., Morgan, R. L., Vine, H., Pedlar, A. & Spencer, R. *Nature* **296**, 331–333 (1982).
- Jones, T. W., Leung, C. M., Gould, R. J. & Stein, W. A. *Astrophys. J.* **212**, 52–59 (1977).
- Martin, P. G. *Mon. Not. R. astr. Soc.* **158**, 63–78 (1972).
- Purcell, E. M. & Spitzer, L. *Astrophys. J.* **167**, 31–62 (1971).
- Purcell, E. M. *Physica* **41**, 100–133 (1969).
- Gold, T. *Nature* **169**, 322 (1952).

Low-level radio flares from Cygnus X-3

Lawrence A. Molnar, Mark J. Reid
& Jonathan E. Grindlay

Center for Astrophysics, 60 Garden Street, Cambridge,
Massachusetts 02138, USA

Cygnus X-3 is a poorly understood but interesting galactic X-ray source. It is unusual in that it shows variable low-state ($S_v < 1$ Jy) radio emission¹ as well as spectacular radio flares ($S_v \sim 20$ Jy). The spectacular flares have long been interpreted as synchrotron emission from an expanding source²; more recent work³ implies that Cyg X-3 may be one of the few identified radio jets in the Galaxy. Flux-density variations with a period of 4.8 h have been discovered at X-ray⁴ and IR⁵ wavelengths, and possibly at γ -ray energies^{6,7} from 10^3 eV to 10^{16} eV. The X-ray emission is probably produced in a binary system composed of a low-mass star orbiting a compact object with an accretion disk and an accretion disk corona⁸. As interstellar absorption⁹ ($A_v = 19$) precludes study of the optical emission, the applicability of this model to Cyg X-3 rests heavily on analogy with other sources with similar X-ray properties (such as 4U1822-37 and 4U2129+47). We report here observations with the Very Large Array (VLA) of Cyg X-3 at several radio wavelengths which show that the radio emission in its low state may be interpreted as arising entirely from the superposition of a series of flares that are wavelength dependent with an apparent period near the 4.8-h X-ray period¹⁰. A model of the source, in which each flare represents the injection of relativistic particles into a small volume which subsequently expands at constant velocity, is qualitatively consistent with the data.

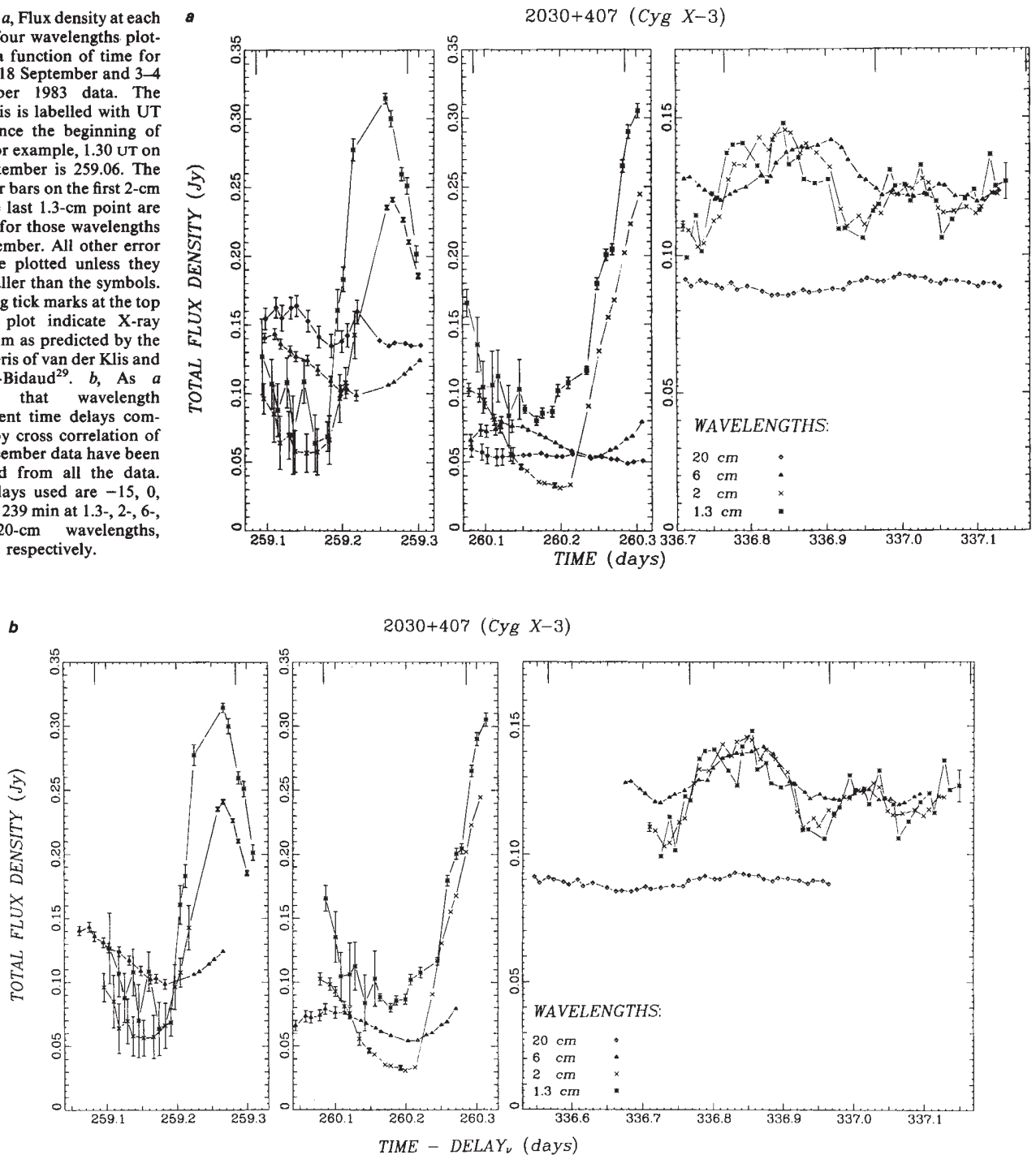
There is little agreement on the location and mechanism of the low-state radio emission. Seaquist and Gregory¹¹ have suggested synchrotron emission from particles embedded in a stellar wind. Vestrand¹² has suggested that γ rays accelerate the relativistic electrons through Compton collisions and pair production. The low flux density of Cyg X-3 in its low state and source confusion make observations difficult, and relatively few observations have been reported. The published data^{13–16} do not give a good definition of either the shape or the time evolution of the radio spectrum.

We have initiated a programme to study the evolution of the radio spectrum of Cyg X-3 in its low state. As part of this programme, we monitored the flux density of Cyg X-3 at the VLA for 6 h from 1.30 to 7.30 UT on both 17 and 18 September 1983 and for 11 h from 16.20 UT on 3 December to 3.20 UT on 4 December 1983 at 1.3-, 2-, 6-, and 20-cm wavelength. Coordinated X-ray observations in September and millimetre observations in December will be reported elsewhere. The flux density at each wavelength is plotted as a function of time in Fig. 1a.

Our procedure was to observe Cyg X-3 at each of the four wavelength bands in succession and then to observe a nearby source (2005+403) at each wavelength for a relative flux density calibration. All Cyg X-3 scans were taken within 15 min of a calibrator scan, which is important at 1.3- and 2-cm wavelengths to eliminate variations of the telescope gain. The flux density of 2005+403 was determined by observations of 3C286. We estimate calibration errors relative to 3C286 to be $< 5\%$ at 1.3 cm and $< 2\%$ at all other wavelengths.

The VLA was in the A configuration¹⁷ in September and in a mixed AB configuration in December. For the first 4 h on 17 September only five inner antennas were available, while for the first 2 h on 18 September only five outer antennas were available. The full array (27 antennas) was available for the rest of the observing time. A 50-MHz bandwidth was used at 1.3-, 2-, and 6-cm wavelengths; a 25-MHz bandwidth was used at 20-cm wavelength to avoid possible interference. Right and left circular polarizations were observed in each of two frequency

Fig. 1 *a*, Flux density at each of the four wavelengths plotted as a function of time for the 17, 18 September and 3–4 December 1983 data. The time axis is labelled with UT days since the beginning of 1983, for example, 1.30 UT on 17 September is 259.06. The 1σ error bars on the first 2-cm and the last 1.3-cm point are typical for those wavelengths in December. All other error bars are plotted unless they are smaller than the symbols. The long tick marks at the top of the plot indicate X-ray minimum as predicted by the ephemeris of van der Klis and Bonnet-Bidaud²⁹. *b*, As *a* except that wavelength dependent time delays computed by cross correlation of the December data have been removed from all the data. The delays used are -15, 0, 52, and 239 min at 1.3-, 2-, 6-, and 20-cm wavelengths, respectively.



bonds. The flux density values in Fig. 1*a* reflect a weighted average of the four channels at each wavelength.

All or part of six flares were observed, two in each observing day. The properties of each flare are such that the longer the wavelength, the lesser the amplitude and the later the time of maximum. Cross-correlation of the flux density versus time for December reveals that the 1.3-cm flux density variations lead the 2-cm variations by 15 ± 5 min, the 2-cm variations lead the 6-cm variations by 52 ± 1 min and that the 2-cm variations lead the 20-cm variations by 239 ± 7 min. The quoted errors are 1σ values determined from the distribution of maximum correlations obtained in 21 simulations in which we added gaussian noise with variance set by the instrumental noise to the raw data. The errors, therefore, also include the effects of the finite time span of the data. Thus the flare that peaks at day 336.83

at 1.3-cm wavelength peaks at day 336.84 at 2 cm, day 336.88 at 6 cm, and at day 337.01 at 20 cm. The 2–1.3 cm and 6–2 cm delays on the two days in September are formally 60% larger than the corresponding December values, although they are not as well determined because we do not have complete coverage of any one flare at all three wavelengths. There is no significant variation in the 20-cm data for September. Figure 1*b* shows the data with the wavelength-dependent delays determined from the December data removed from all the data. It reveals how well a single set of delays aligns the maxima and minima at all wavelengths.

We used a correlation analysis to estimate the time between subsequent flares on the three days. Cross correlation of the 1.3-cm and 6-cm flux densities on 17 September yielded significant maxima at -0.147 and $+0.079$ days, implying a

2030+407 (Cyg X-3)

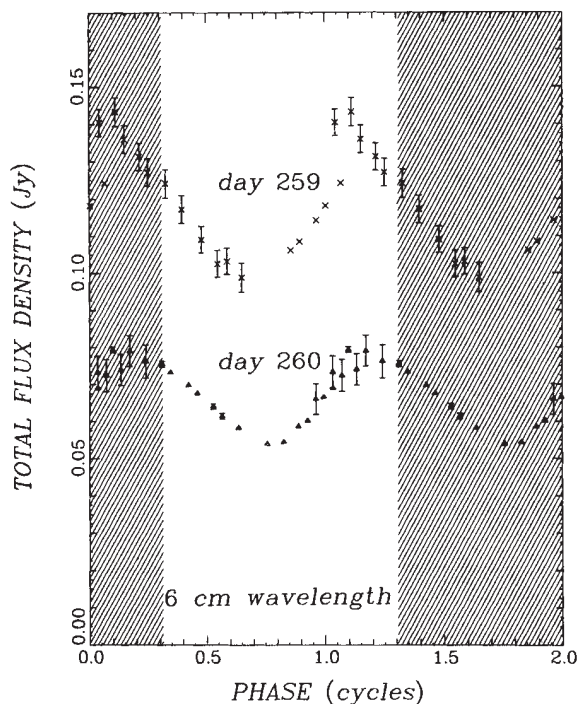


Fig. 2 Flux density at 6-cm wavelength from 17 September ($\times$) and 18 September (Δ) plotted against X-ray phase with zero phase corresponding to X-ray minimum. While we have folded the data on the X-ray period, similar graphs would result using periods in the range 4.8–5.1 h. Each datum is plotted twice for clarity, once in the shaded portion of the plot.

separation of 5.4 h between the two flares on that day. Cross-correlation of the 2-cm and 6-cm flux densities on 18 September yielded significant maxima at -0.167 and $+0.056$ days, implying a separation of 5.4 h. Autocorrelation of the 1.3-cm and 2-cm December data yielded a maximum at $+0.207$ days, implying a flare separation of 5.0 h. The September estimates are limited by the incomplete coverage of the flares and may have biases as large as 0.5 h, while the December estimate is limited by the instrumental noise (1σ is 0.1 h). The similar time separations and the relatively smooth light curves (see Figs 1, 2) suggest that the flares may be periodic. Assuming periodicity, the observed separations require that the 2-cm flare at day 260.30 is the fifth one following the 2-cm flare at day 259.26. With this information, cross-correlation between the 2 days of the same wavelength data yields 5.06, 4.98, and 4.82 h at 1.3-, 2-, and 6-cm wavelengths, respectively. This is consistent with the 4.8-h X-ray period although favouring a value about 0.1 h longer.

We estimated the probability that the source is not periodic at the X-ray period with the following χ^2 test. We binned the data for the two September days at the known X-ray period with bin sizes corresponding to the sampling interval (15 min). We computed a χ^2 value using the variance of the points within the bins to predict the scatter among the bin-averaged fluxes. The resulting probabilities that the deviations observed could have occurred by chance are 14.7%, 0.3%, $<0.01\%$, and 97.3% for 1.3-, 2-, 6-, and 20-cm wavelengths, respectively. (For a conservative estimate of χ^2 , we used a value of the variance within the bins equal to the nominal value plus the 1σ error on that value¹⁸.) Thus the flux densities at 2- and 6-cm wavelengths are probably periodic. Figure 2 shows the 6-cm data from September folded at the X-ray period, which graphically illustrates the apparent periodicity.

Using this same χ^2 analysis, we searched periods from 1.7 to

5.8 h, the range in which enough data exists to perform the test. The peak values occurred at 5.0, 4.9, and 4.8 h, for 1.3-, 2-, and 6-cm wavelengths, respectively. The only other significant probability was at 6 cm at 4.0 h ($<0.01\%$), an alias of the 4.8-h period arising from the gap in sampling.

Although the present data set cannot make a significant distinction among periods in the range 4.8–5.1 h, the results of the correlation analysis and the χ^2 binning analysis point to a period slightly longer than 4.8 h. Furthermore, the phase of the radio flare at the time of X-ray minimum in the December data differs from that in the September data, also consistent with a small difference between the radio and X-ray periods. An alternative hypothesis is that the radio emission is periodic at the X-ray period with some phase jitter. Observations over a greater time baseline are needed to confirm the periodicity and to establish any difference between the radio and X-ray periods.

A model in which each flare results from the injection of relativistic particles emitting synchrotron radiation into a small volume which subsequently expands at constant velocity^{19–21} is qualitatively consistent with the data. The 1 m arc s size measured by Geldzahler *et al.*¹⁶ implies a brightness temperature $\sim 10^9$ K, supporting the hypothesis that the emission is nonthermal. The expanding source model assumes magnetic flux conservation and that adiabatic energy losses dominate synchrotron and inverse Compton losses. At early times in a given flare, synchrotron self-absorption is important and the total flux density grows with the increasing source solid angle (and to a lesser extent with the decreasing magnetic field strength). After the source becomes optically thin, the flux density declines due to the decreasing magnetic field strength and adiabatic losses. The peak in the spectrum, separating the two optical depth regimes, moves to longer wavelengths and lower flux densities with time—characteristics seen in the Cyg X-3 data. The rate at which the spectral peak moves is anticorrelated with the magnitude of the flare; this is also seen in the Cyg X-3 data as longer delays between flux-density variations at the different wavelengths in the September than in the December flares. If there are periodic particle injections, subsequent flares are only cleanly separated in time at the shortest wavelengths, making observations at these wavelengths ($\lambda < 2$ cm for Cyg X-3) important for establishing an optically-thin spectral index for a single flare.

The model incorporates three basic observational parameters: the peak flux density at a reference wavelength, the time required to reach peak flux density at that wavelength, and the optically-thin spectral index. The photon spectral index is directly related to the electron-energy spectral index. The remaining two observational parameters provide two relations between three physical parameters: the magnetic field, the particle density (each at some reference size or time) and the expansion velocity. Thus, we need one further constraint to determine the model completely.

Were the flare to begin at an arbitrarily small size, the magnetic field strength and the photon energy density would be arbitrarily large. Hence, the model is only applicable at sizes greater than some minimum size at which the model assumptions for energy losses are valid. Particles injected into a smaller volume with the physical parameters computed from the model would rapidly lose their energy. (The parameter values for the minimum size are similar to those obtained by assuming equipartition of energy or minimum energy.) For the second 17 September flare, for example, this size is $\sim 6 \times 10^{12}$ cm, corresponding to 0.07 m arc s at a distance^{22,23} of 12 kpc. An upper limit on this size comes from requiring that the observed time scale of variation is less than a light crossing time. The 1.3-cm flux density increased by a factor of 4 over 3,000 s in the September flare. Hence a conservative upper limit on the source radius at the time of particle injection is 10^{14} cm.

This upper limit on the source size combined with a reasonable upper limit on the total energy (10^{43} ergs per flare) restricts the range of expansion velocities to $0.05c$ – $0.25c$, which would correspond to the source diameter increasing at 0.06 – 0.3 m arc s h^{-1} .

Assuming equipartition of particle and magnetic-field energies implies a value of the magnetic field (at source radius 10^{13} cm) of 50 G and a relativistic electron energy density of 10^2 erg cm $^{-3}$. The limits on source size and total energy restrict these quantities to within two orders of magnitude of their equipartition values.

More detailed analysis of the current data (including radiative-energy loss mechanisms, light crossing time, and free-free absorption by thermal particles) will allow a more stringent test of the model and a better constraint on the parameters. The velocity range and the minimum source size predict that the expansion of this source may be directly measurable with very long baseline interferometry (VLBI) techniques. By measuring the expansion velocity, VLBI observations would completely determine the physical parameters of the model.

If the apparent radio periodicity is supported by additional observations, the synchrotron emission region of Cyg X-3 is undergoing periodic particle injection. Periodic nonthermal radio emission has been observed from the much longer period binaries Cir X-1 (ref. 24), LS I +61°303 (ref. 25), and SS 433 (refs 26, 27), although the flaring mechanisms in those more widely separated systems are probably different. In the case of Cyg X-3, periodic production of relativistic particles may be a natural result of a surge in mass transfer (as suspected^{26,27} for SS 433) at periastron passage in an eccentric orbit. The possible ~0.1-h difference between the radio and X-ray periods could arise from the beat between the 4.8-h orbital period and the possible 19-day period found from long-term timing studies of Cyg X-3 (ref. 28). The 19-day period is, in fact, consistent with apsidal motion of a binary orbit with an orbital eccentricity²⁸ of 0.03. A precise value for the radio period would, therefore, provide important information on the dynamics of the binary system itself.

The minimum size derived from our model is $60(M/M_{\odot})^{-1/3}$ binary radii, where M is the total mass of the system and $M_{\odot}$ is the mass of the Sun. Hence the expanding synchrotron source model requires a mechanism for transporting energy from the binary system to the electrons in the outlying radio emission region. Vestrand's²² suggestion of γ -ray interactions with outlying thermal material is possible, although bulk transport of energy in jets (as in SS 433) is perhaps more promising.

The VLA is a facility of the National Radio Astronomy Observatory, which is operated by Associated Universities, Inc., under contract with the NSF.

Variation in observed coronal calcium abundance of X-ray flare plasmas

J. Sylwester*, J. R. Lemen† & R. Mewe‡

* Space Research Centre, Polish Academy of Sciences, Kopernika 11, Wrocław, Poland

† Mullard Space Science Laboratory, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

‡ Laboratory for Space Research, Beneluxlaan 21, Utrecht, The Netherlands

Variations in chemical composition during solar flares have been inferred from elemental abundance changes in cosmic ray fluxes, but have so far not been detected spectroscopically. We present here the first spectroscopic evidence for the variation of the coronal calcium abundance in high-temperature solar flare plasmas. The analysed data consist of the high-resolution X-ray flare spectra ($\lambda/\Delta\lambda \approx 4,000$) observed with the Bent Crystal Spectrometer (BCS) on board the Solar Maximum Mission (SMM) satellite and described in detail by Acton *et al.*¹. The observed abundance variation has important consequences for the analysis and interpretation of XUV and X-ray spectra.

In its lowest energy channel the BCS observes the X-ray spectrum in the vicinity of the helium-like resonance, forbidden and intercombination emission lines of Ca XIX and the continuum, to the blue of the resonance line (see Fig. 1b in ref. 2). From BCS spectra we derive the line-to-continuum ratio (I_L/I_C) for the resonance line of Ca XIX ($\lambda = 3.1781$ Å) as a function of temperature.

We express the I_L/I_C flux ratio as a function of electron temperature (T) for an optically-thin, isothermal plasma as:

$$\frac{I_L}{I_C} \propto A_{Ca} \frac{N_{CaXIX}}{N_{Ca}} \frac{\Omega(T)}{G(T)} \quad (1)$$

where A_{Ca} is the calcium abundance relative to hydrogen, $\Omega(T)$ is the effective collision strength for the line emission (see ref. 3), and $G(T)$ is the generalized Gaunt factor describing the continuum emission (see ref. 4). N_{CaXIX}/N_{Ca} , the concentration of the helium-like calcium ion relative to the total number of calcium ions, also depends on temperature.

An example of the observed I_L/I_C evolutionary time behaviour during a flare is shown in Fig. 1 for the flare which occurred on 14 July 1980 at 08.25 UT. The electron temperature has been estimated from the ratio of the dielectronic satellite Ca XVIII $^2P_{1/2}-^2D_{3/2}$ ($\lambda = 3.208$ Å, line k in the notation of Gabriel⁵) intensity to the Ca XIX resonance line ($^1S_0-^1P_1$) intensity using the atomic data given by Bely-Dubau *et al.*³. This line ratio is a sensitive measure of the electron temperature and because the satellite is formed by dielectronic recombination, the accuracy does not depend on the assumed ionization balance calculation. The indicated $\pm 1\sigma$ error bars were estimated from the counting statistics of the resonance and satellite lines.

The observed I_L/I_C evolution during the heating phase (*ab* in Fig. 1) shows a hysteresis behaviour with temperature that is not predicted by equation (1) which is single-valued for all values of temperature. The hysteresis of I_L/I_C with temperature has been observed in several flares and three possible explanations are being investigated: the presence of a small amount (a few per cent of the total emission measure) of very hot plasma ($T > 10^8$ K); emission produced by non-thermal electrons; or a variation of the calcium abundance during the heating phase. Although a transiently-ionizing plasma could possibly give rise to the hysteresis of I_L/I_C with temperature, the time of ≤ 10 s to reach equilibrium for a typical density of 10^{10} cm $^{-3}$ is much shorter than would be necessary to explain the observations.

Figure 2 shows the cooling phases (*bc* in Fig. 1) for the 14 July 1980 flare and for the flare which occurred on 29 June 1980 at 18.00 UT. During the cooling phase the temperature behaviour

Received 14 March; accepted 26 June 1984.

- Braes, L. L. E. & Miley, G. K. *Nature* **237**, 506 (1972).
- Gregory, P. C. *et al. Nature phys. Sci.* **239**, 114-117 (1972).
- Geldzahler, B. J. *et al. Astrophys. J. Lett.* **273**, L65-L69 (1983).
- Brinkman, A. *et al. IAU Circ. No.* 2446 (1972).
- Becklin, E. E. *et al. Nature* **245**, 302-304 (1973).
- Lamb, R. C., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Thompson, D. J. *Astrophys. J. Lett.* **212**, L63-L66 (1977).
- Samorski, M. & Stamm, W. *Astrophys. J. Lett.* **268**, L17-L21 (1983).
- White, N. E. & Holt, S. S. *Astrophys. J.* **257**, 318-337 (1982).
- Weekes, T. C. & Geary, J. C. *Publ. astr. Soc. Pacif.* **94**, 708-712 (1982).
- Molnar, L. A., Reid, M. J. & Grindlay, J. E. *IAU Circ. No.* 3885 (1983).
- Sequist, E. R. & Gregory, P. C. *Astrophys. Lett.* **18**, 65-68 (1977).
- Vestrand, W. T. *Astrophys. J.* **271**, 304-312 (1983).
- Hjellming, R. M., Hermann, M. & Webster, E. *Nature* **237**, 507-508 (1972).
- Hjellming, R. M. & Balick, B. *Nature* **239**, 443-446 (1972).
- Mason, K. O. *et al. Astrophys. J.* **207**, 78-87 (1976).
- Geldzahler, B. J., Kellermann, K. I. & Shaffer, D. B. *Astr. J.* **84**, 186-188 (1979).
- Thompson, A. R., Clark, B. G., Wade, C. M. & Napier, P. J. *Astrophys. J. Suppl.* **44**, 151-167 (1980).
- Danese, L., DeZotti, G. & diTullio, G. *Astr. Astrophys.* **82**, 322-327 (1980).
- Shklovskii, I. S. *Soviet. Astr.* **9**, 22-23 (1965).
- van der Laan, H. *Nature* **211**, 1131-1133 (1966).
- Kellermann, K. I. & Pauliny-Toth, I. I. K. A. *Rev. Astr. Astrophys.* **6**, 417-448 (1968).
- Dickey, J. M. *Astrophys. J. Lett.* **273**, L71-L73 (1983).
- Cawley, M. F., Gibbs, K. & Weekes, T. C. *Proc. 18th int. Cosmic Ray Conf.* (in the press).
- Nicolson, G. D., Feast, M. W. & Glass, I. S. *Mon. Not. R. astr. Soc.* **191**, 293-299 (1980).
- Taylor, A. R. & Gregory, P. C. *Astrophys. J.* **255**, 210-216 (1982).
- Grindlay, J. E. *et al. Astrophys. J.* **277**, 286-295 (1984).
- Band, D. L. & Grindlay, J. E. *Astrophys. J.* (submitted).
- Bonnet-Bidaud, J. M. & van der Klis, M. *Astr. Astrophys.* **101**, 299-304 (1981).
- van der Klis, M. & Bonnet-Bidaud, J. M. *Astr. Astrophys.* **95**, L5-L7 (1981).

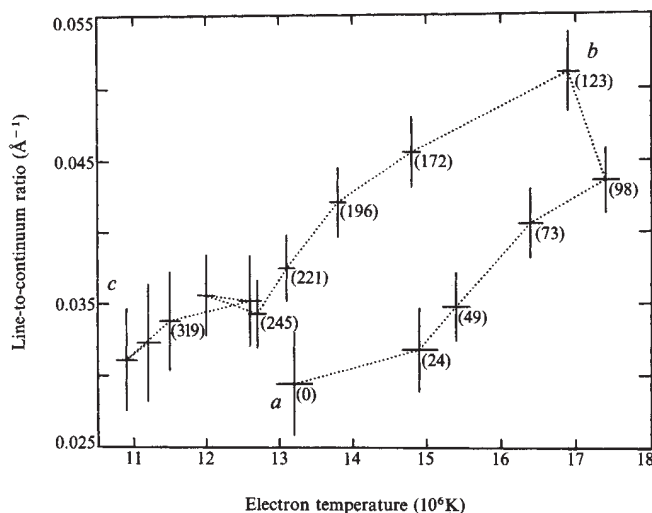


Fig. 1 The line-to-continuum ratio (I_L/I_C) evolutionary time behaviour for the flare on 14 July 1980. Times in seconds after 8.24.35 UT are indicated in parentheses. The mean electron temperature has been estimated from the satellite-to-resonance line ratio (line k to line w in the notation of Gabriel²) using the atomic data given by Bely-Dubau *et al.*³. *ab* corresponds to the heating phase, *bc* to the cooling phase; $\pm 1\sigma$ error bars are indicated for the temperature estimations and $\pm 3\sigma$ error bars are indicated for the I_L/I_C values.

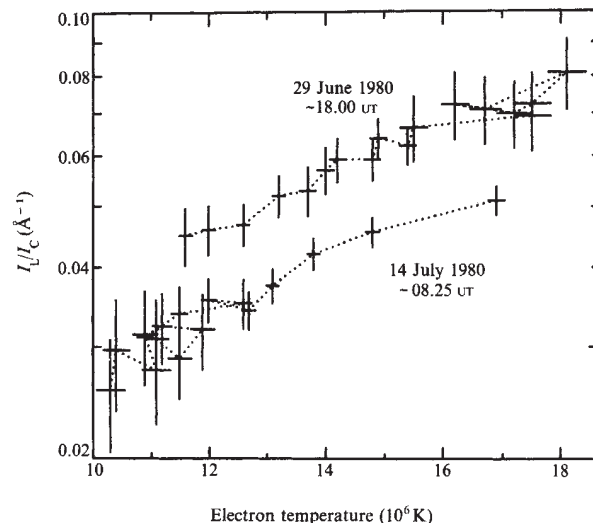


Fig. 2 I_L/I_C for cooling phases of two different flares. The error bars are as in Fig. 1. The constant vertical shift corresponds to a variation in the calcium abundance by a factor of 1.4.

of I_L/I_C had the same form for 13 flares we considered, except that the normalization factor varied by as much as 2.5 between flares in extreme cases. This observed variation in the normalization factor is probably due to the different coronal calcium abundance in each flare. We have ruled out multi-thermal effects as the cause of the observed variation in the normalization of I_L/I_C during the cooling phase. As all the factors in equation (1) apart from the abundance depend on atomic physics and are thus not expected to vary, we are left to conclude that we have observed for the first time a flare-to-flare variation of the coronal calcium abundance.

The BCS continuum fluxes were checked by comparison with observations obtained with the broad-band Hard X-Ray Imaging Spectrometer⁶ (HXIS), on the SMM spacecraft. Preliminary work carried out on HXIS data has shown a similar hysteresis behaviour with temperature for the ratio of the measured-to-predicted count rate for channel 2 (5.5–8.0 keV). In this energy band there is a contribution of ~30–50% from Fe $K\alpha$ line radiation and thus the observed behaviour may be attributable to variations in the iron elemental abundance. If proven, these observations could provide spatial information concerning the elemental composition distribution.

Following theoretical considerations^{7,8} that thermal diffusion could have a significant role in differentiating the elemental abundance structure inside the Sun, several authors have investigated theoretically the role of this effect on the composition of the upper solar atmosphere. It has been predicted^{9–11} that a layer should exist within the transition region where the heavy elemental abundances (relative to hydrogen) should be increased by a significant factor—sometimes much greater than 10—but to date there has been no observational evidence^{12,13} confirming these results. The only positive detection of varying elemental composition of the solar-originated plasma has been found in solar wind composition measurements¹⁴ (depletion of helium, enrichment of heavier elements in the post-flare-associated shocks) and in observations of the middle-energy, solar cosmic rays¹⁵ (enrichment of heavier elements proportional to atomic weight).

The observed variation (see Fig. 2) in the calcium abundance for different flare plasmas could be attributed to: (1) variations in the elemental composition of the pre-flare plasma or (2) variations in the speed of the elemental separation process

during flare heating. The fundamental process providing the basis for any elemental separation is probably differential diffusion occurring across steep temperature and pressure gradients. Detailed considerations of the role of diffusion in cases (1) and (2) will be given elsewhere.

The calcium abundance variation probably represents the first spectroscopic observation of a change in the flare plasma chemical composition. This discovery has important consequences for the analysis of flare spectra in the X-ray region and at longer wavelengths. Many diagnostic techniques rely on the assumption that the abundance of a given element is constant or that the relative chemical composition is constant in time and space in the solar corona. This assumption will now have to be reconsidered. For example, differential emission measure analysis may be affected, as well as temperature and density line ratio diagnostics which rely on lines belonging to different elements. It is possible that the relative chemical composition of heavier elements is not altered significantly from flare to flare, but this needs to be verified experimentally.

The potential of abundance variation analysis as a method of transition-region diagnostics will probably stimulate future observing programmes. The chemical composition should be considered as an additional parameter for characterizing the flare plasma.

We thank Professors J. L. Culhane, J. Jakimiec and C. deJager for helpful discussions. The XRP experiment is a collaborative programme between Lockheed Palo Alto Research Laboratories, the Rutherford Appleton Laboratory and the Mullard Space Science Laboratory. J.S. acknowledges financial support from the SERC during his visits to Mullard Space Science Laboratory and R.M. acknowledges financial support from the Space Research Organization of the Netherlands (SRON).

Received 19 March; accepted 20 June 1984.

1. Acton, L. W. *et al. Solar Phys.* **65**, 53–71 (1980).
2. Culhane, J. L. *et al. Astrophys. J. Lett.* **244**, L141–L145 (1981).
3. Bely-Dubau, F. *et al. Mon. Not. R. astr. Soc.* **201**, 1151–1169 (1982).
4. Mewe, R. & Gronenschild, E. H. B. M. *Astrophys. Suppl. Ser.* **45**, 11–52 (1981).
5. Gabriel, A. H. *Mon. Not. R. astr. Soc.* **160**, 99–119 (1972).
6. van Beek, H. F., Hoyng, P., Lafleur, B. & Simnett, G. M. *Solar Phys.* **65**, 39–52 (1980).
7. Chapman, S. & Cowling, T. G. *Mathematical Theory of Non-uniform Gases* (Cambridge University Press, 1952).
8. Aller, L. H. & Chapman, S. *Astrophys. J.* **132**, 461–472 (1960).
9. Delache, P. *Annls Astrophys.* **30**, 827–860 (1967).
10. Nakada, M. P. *Solar Phys.* **7**, 302–320 (1967).
11. Tworowski, A. S. *Astrophys. Lett.* **17**, 27–30 (1975).
12. Mariska, J. T. *Astrophys. Lett.* **235**, 268–273 (1980).
13. Dere, K. P. & Mason, H. E. *Solar Active Regions—Monogr. Skylab Solar Workshop III* (ed. Orrall, F. Q.) 129–164 (Colorado Associated University Press, 1981).
14. Geiss, J. *Space Sci. Rev.* **33**, 201–217 (1982).
15. Fan, C. Y., Gloeckler, G. & Hovestadt, D. *Solar Gamma-, X- and EUV Radiation* (ed. Kane, S. R.) 411–421 (Reidel, Dordrecht, 1975).

Snow and ice feedbacks prolong effects of nuclear winter

Alan Robock

Department of Meteorology, University of Maryland, College Park, Maryland 20742, USA

Recent studies using climate models¹⁻⁶ have suggested that drastic surface cooling (the "nuclear winter"¹) caused by smoke and dust would follow a large-scale nuclear war, with possible drastic effects on the biosphere⁷. None of these studies looked at the long-term effects of time of year on the results. Moreover, although the general circulation model experiments of Covey *et al.*⁴ had snow/albedo feedback on land, none of the experiments considered long-term seasonal cryospheric interactions with the forcing, especially the sea ice/thermal inertia feedback⁸. Although, in the previous studies, the presumed transport of smoke and dust to the Southern Hemisphere was mentioned, no detailed investigation of its effects has been made. I investigate here the interactions of snow and sea ice with the forcing, and look at the effects over a period of several years for a variety of strengths of the forcing and latitudinal distribution of the nuclear smoke and dust, including Southern Hemisphere injection. The seasonal dependence of both the forcing and the response are also investigated. Previous results are supported and possible longer lasting effects are revealed. The considerable uncertainties that remain are discussed.

The experiments were conducted with a seasonal energy balance climate model based on that of Sellers^{9,10}. This model, described in detail elsewhere⁸, has been used to simulate the effects of volcanic eruptions¹¹⁻¹⁴ and several other forcings^{14,15} on climate. It uses 15-day time steps and solves for surface air temperature over land and water separately, in 10° latitude bands. Incoming and outgoing radiation are considered in detail and horizontal energy transports by the atmosphere and ocean are parameterized. Although energy balance climate models successfully reproduce mean properties of the observed climate system, details which may be important for the nuclear winter problem cannot be considered explicitly. Variations of atmospheric lapse rate, such as surface inversions, cannot be modelled. The three-dimensional circulation of the atmosphere, and its response to, and interaction with, the smoke and dust must be externally specified in this experiment, rather than internally calculated. The simple ocean model, with fixed mixed layer depth (from 90 m at the Equator to 40 m at the poles) and diffusive heat transport, prevents detailed consideration of ocean dynamical response. This climate model has the advantage, however, that it can be run rapidly on the computer and considers snow and ice feedbacks explicitly. The experiments reported here must, therefore, be interpreted as indicative of possible long-term interactions with the climate system that may result from forcing by nuclear smoke and dust. The exact surface-temperature responses would be different if the above details were included in a more complex model.

This climate model represents an energy balance for the vertically integrated surface-atmosphere system. In this experiment, the model is used to simulate the climate system beneath the nuclear smoke and dust layer. Turco *et al.*'s calculations¹ of the solar-energy flux at the ground for different scenarios were used to force the model. The scenarios of Turco *et al.* were first extrapolated to 400 days where the reduction of solar radiation went to zero for all cases. To force the model, the solar flux at each latitude band where there was smoke and dust was reduced by the same percentage for the particular scenario as given in their Fig. 4. Because Turco *et al.*'s values were for smoke and dust spread evenly over the Northern Hemisphere, they were adjusted according to the formula:

$$T' = T^f$$

where T' is the per cent transmission of solar radiation actually

Table 1 Forcing variations of the experiments

Time dependence of the forcing

Reduction of insolation according to Turco *et al.*¹ (their Fig. 4)

Case	
1	5,000 MT-baseline
14	100 MT-city attack
17	10,000 MT-severe

Latitudinal dependence of forcing

		f
Ref. 4	30–70° N	2.273
Ref. 3	10–90° N	1.205
Global	90° S–90° N	0.500
Gradual Southern Hemisphere (SH)	10–90° N	1.205
	0–90° N	0.667
	0–90° S (after 3 months)	0.333

Time war begins

	Control case
Winter (1 January)	Ref. 4 (30–70° N), $f = 2.273$
Spring (1 April)	Ref. 1 5,000 MT-baseline (case 1)
Summer (1 July)	Summer (war begins 1 July)
Autumn (1 October)	

used as forcing, T is the per cent transmission from ref. 1, and f is a latitudinal factor equal to the reciprocal of the percentage of the Earth covered by the smoke in units of Northern Hemisphere area. The equation was derived by considering that solar flux reduction is due to an exponential decrease in radiation for a linear increase in smoke amount. The values of f are given in Table 1 for the different cases used here.

All other components of the climate model were not changed, including the outgoing long-wave radiation. This way of forcing the model implies that the smoke and dust layer is high in the atmosphere where it has little impact on the long-wave radiation, due to its particle size distribution. The long lifetime of the smoke in the upper troposphere follows from the results of MacCracken² and Turco *et al.*¹ that convection is essentially cut off due to the extreme solar absorption by the smoke. The results that follow depend on these assumptions.

Table 1 also summarizes the forcings used in these experiments. Three of Turco *et al.*'s scenarios were used for the time dependence of the forcing, case 1 (5,000 MT-baseline), case 14 (100 MT-city attack) and case 17 (10,000 MT-severe). These scenarios do not necessarily represent the most realistic or most probable radiative forcings, but were chosen to allow the climate model to assess the importance of feedbacks on the long-term response. The effects of the time of initiation of the war were tested by starting the forcing in four different seasons. Four different latitudinal distributions were used, two to correspond to previous GCM experiments, from ref. 4 (30° N–70° N uniformly distributed) and from ref. 3 (10° N–90° N uniformly distributed), a globally uniform case as an extreme limit, and a case called Gradual Southern Hemisphere where the smoke spreads gradually to the Southern Hemisphere after 3 months. This last case tests the suggestions of previous models that a cross-equatorial circulation would develop due to the large temperature gradient caused by the smoke which would then transport the smoke to the Southern Hemisphere even if no bombs were exploded there.

Because of the many different possible combinations of forcings that could have been used, it was decided to perform a control case, and then change one factor at a time from this. The control case chosen used Turco *et al.*'s 5,000 MT-baseline scenario (their case 1) and Covey *et al.*'s⁴ latitude distribution (30° N–70° N), with the war beginning in summer.

Nine different runs were made. They are listed in Table 2 and the changes in surface temperature from the pre-war values are plotted in Fig. 1 as a function of latitude and time of year. Only the first 4 yr of results are shown. After this time, the temperature response shows a slow oscillation with the same pattern as in year 4 as the effect gradually diminishes. The same behaviour

Table 2 List of results shown in Fig. 1 and maximum cooling produced

Run	Latitudinal dependence	Time dependence	Time war begins	Results shown for:	Maximum temperature change (°C)
1a	Ref. 4	Ref. 1 Case 1 (baseline)	Summer	Land (control)	-21.4
1b	Ref. 4	Ref. 1 Case 1	Summer	Ocean	-10.7
1c	Ref. 4	Ref. 1 Case 1	Summer	Zonal average	-15.4
2	Ref. 4	Ref. 1 Case 1	Autumn	Land	-8.3
3	Ref. 4	Ref. 1 Case 1	Winter	Land	-8.1
4	Ref. 4	Ref. 1 Case 1	Spring	Land	-16.9
5	Ref. 3	Ref. 1 Case 1	Summer	Land	-21.1
6	Gradual Southern Hemisphere (SH)	Ref. 1 Case 1	Summer	Land	-19.4
7	Global	Ref. 1 Case 1	Summer	Land	-18.3
8	Ref. 4	Ref. 1 Case 17 (10,000 MT-severe)	Summer	Land	-22.9 (year 2) -16.4
9	Ref. 4	Ref. 1 Case 14 (100 MT-city)	Summer	Land	-19.3

can be seen in response to forcing from volcanic eruptions both for a 9-yr climate model simulation^{11,12} and in surface temperature observations for the past 90 yr (ref. 16).

Results 1a, 1b and 1c (Table 2) are all from the same control run, and present the surface-air temperature change for land grid areas, ocean grid areas and the zonal average. The effect is initially much larger over land, due to the lower thermal inertia. The maximum cooling is 21.4°C and occurs in the 50–60° N latitude band 45–60 days after the beginning of the war. The Northern Hemisphere average land cooling at this time is 11.5°C, which agrees well with MacCracken². Over the ocean grid boxes, the maximum effect is delayed and smaller, but occurs in the northernmost grid box, due to lower thermal inertia here caused by the presence of sea ice, and has a value of 10.7°C 90–105 days after the war starts. Even in the mid-latitude ocean areas, large temperature drops are found, which have not been found in any of the previous studies. A much more detailed ocean model than the mixed layer with fixed depth and diffusive heat transport used here is needed to investigate these effects further. The zonal average maximum effect of 15.4°C agrees almost exactly with MacCracken's results, as does the land value with the other previous results.

The climate model response in the first year after the war begins can be easily understood as a linear response to decreased insolation. During the second year, as the forcing disappears, the nonlinear effects of the snow and ice feedbacks become apparent. The large area of cooling of more than 5°C found in the mid and high latitudes in the summer over land is due to the snow/albedo feedback. The additional snow produced due to the large previous cooling reflects more solar radiation, enhancing the cooling. This feedback is only important in the summer when there is substantial insolation⁸. Over the ocean, the largest cooling in the second year is at the poles in the winter. This pattern, produced by the sea ice/thermal inertia feedback⁸, is caused by the enhanced amplitude of the seasonal cycle of temperature. This occurs when cooling produces more sea ice which lowers the thermal inertia of the ocean. The sea ice/albedo feedback, which in the absence of the thermal inertia feedback would produce the same pattern that the snow/albedo feedback produces on land, with enhanced sensitivity in the summer, is completely overwhelmed by the much stronger thermal inertia feedback over the ocean. (The feedbacks discussed here ignore possible effects of smoke particles settling on the snow and ice and hence lowering the albedo until the next snowfall¹⁷. These dirty snow effects may weaken, or even enhance, these feedbacks.)

The mixture of these two feedbacks from land and ocean can be seen in the pattern of the zonal average for the second year. The overall pattern over land begins to resemble the ocean pattern at the end of the second year, as the ocean effects dominate due to the horizontal mixing. In the third year, only a small area of sensitivity is evident in the summer high latitudes over land from the snow/albedo feedback. By the end of the

second year, 18 months after the start of the war, cooling over the oceans is of the same amplitude as over land, and after that it is even larger, as the ocean thermal inertia slows the recovery to pre-war temperatures. Note that these feedbacks and patterns are also evident in the equilibrium response of the climate model to external forcing, such as changing solar constant or CO₂ (refs 8, 18), as opposed to the transient response shown here.

Because of the cryosphere feedbacks discussed above, the response of the climate to nuclear war is longer and larger than previously found by Turco *et al.*¹ who did not include these feedbacks. It would presumably be very difficult to grow crops not only in the summer of the year that the war began, but also in the summer of the next year, with temperatures over land almost 6°C (>10°F) colder than normal.

Several other experiments were conducted to test the effects of time of year that the war begins, latitudinal spread of the smoke, and strength of the forcing. To save space, only the response over land is shown for comparison with the control case. The response over land is probably of more interest anyway, because it is larger than the first year response over the sea, and humans and agriculture are affected. The relative effects over the ocean, as shown for the control, are the same for all these cases.

Runs 2, 3 and 4 tested the effects of beginning the war in autumn, winter and spring. As can be seen in Fig. 1, the effects in the autumn and winter cases are not as large as in the control. This is because there is less insolation at the time of the maximum atmospheric loading from smoke and dust, and, therefore, less absolute reduction of incoming energy. In the autumn case, the initial effect is larger than the largest effect for the winter case, but it occurs in the autumn and does not last as long. The cooling during the following summer, however, is more than 4°C during the entire summer from 40° N to the Pole. In the winter case, although the initial effect is small, enough smoke remains to cause a large effect the following summer, with a cooling as large as 8°C. This effect was not seen by Covey *et al.*⁴ because they ran their winter experiment for only 20 days. The spring case is quite similar to the summer case with large cooling during the first summer, and substantial, although not quite as large, cooling during the second summer.

Runs 5, 6 and 7 illustrate the effects of spreading the nuclear smoke and dust over a greater latitudinal extent. In the Aleksandrov-Stenchikov³ case (run 5), the results are almost the same as the control case, but the cooling extends over more latitudes and persists at a slightly larger amplitude in the years after the war. Although the smoke concentration is lower in this case than in the control, there is still so much smoke that almost all of the sunlight is initially prevented from reaching the ground. (The small regions of enhanced or decreased sensitivity found in the tropics in this and the following runs are due to a small model instability relating to its inability to calculate the circulation correctly under such extreme forcing. They are not real, and a slight smoothing would produce a more realistic response.)

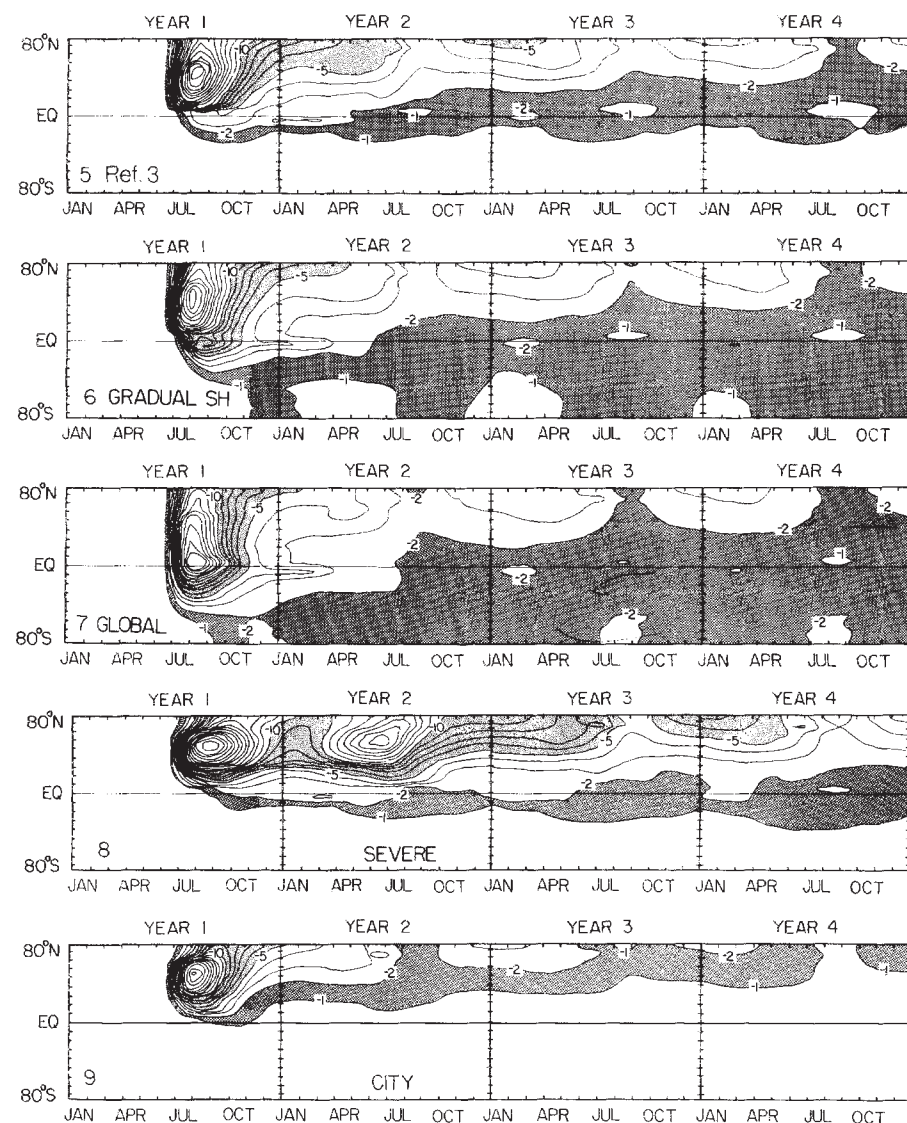
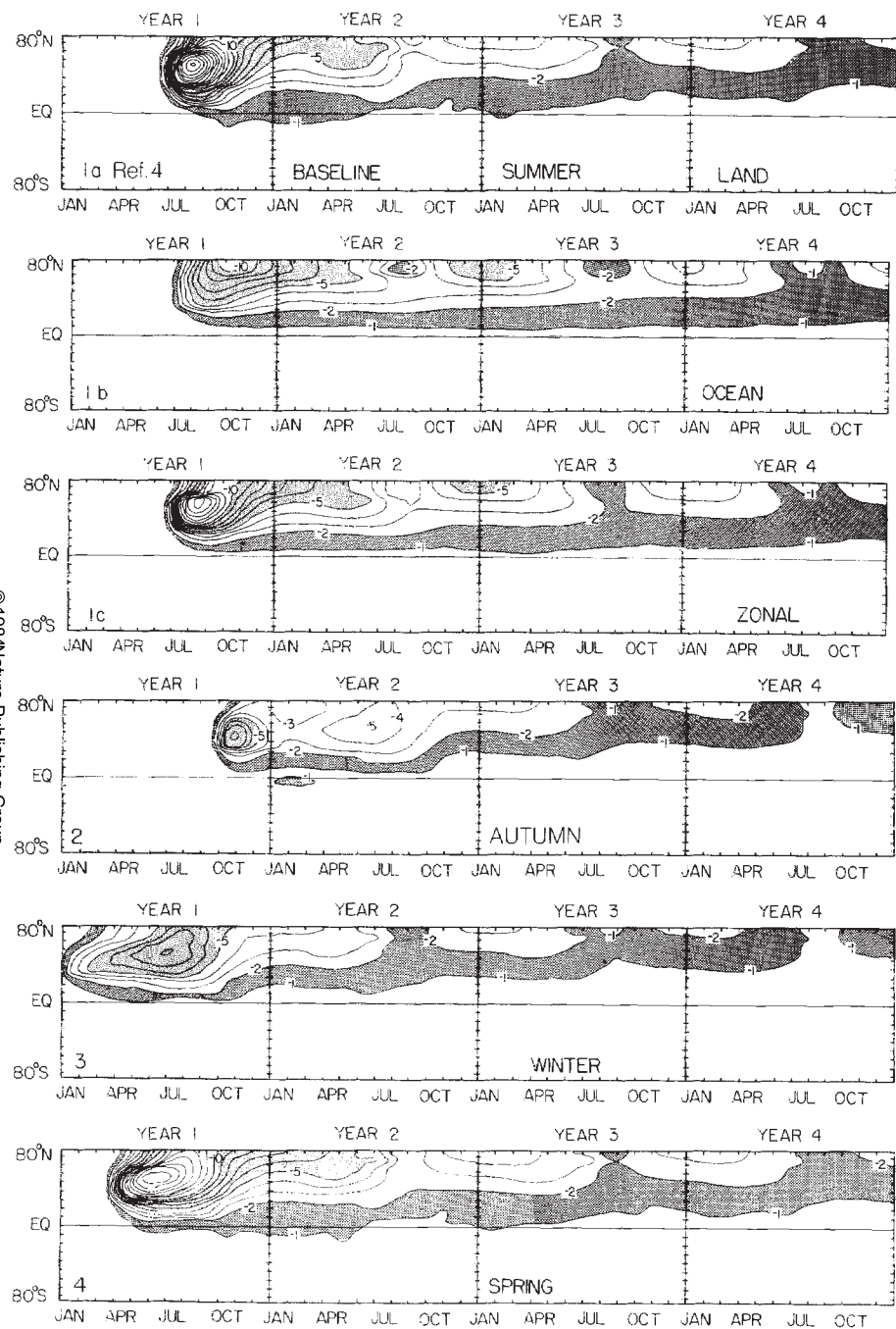


Fig. 1 Temperature change (nuclear war experiments minus unperturbed climate) as function of latitude and time for the runs listed in Table 2. Contours are every 1 °C starting at -1 °C. Temperature drops between -1 °C and -2 °C, and between -5 °C and -10 °C, are shaded.

In the Gradual Southern Hemisphere case (run 6), the Northern Hemisphere response is again about the same as in the control. Although the response is slightly less in year 2, it is larger in years 3 and 4 because the forcing at lower latitudes makes the effects last longer through ocean cooling. In this case, we see maximum cooling of about 8 °C in the tropics, and much smaller cooling in the Southern Hemisphere. The smaller Southern Hemisphere effect is caused by the much larger thermal inertia due to the much larger percentage of oceans in this hemisphere, as well as the lack of land in the high mid-latitudes on which to produce a snow/albedo feedback. This effect is illustrated in run 7, where the forcing is the same at all latitudes, but the response is much less in the Southern Hemisphere. This is partly because the forcing started in Southern Hemisphere winter. A global run starting in Southern Hemisphere summer (not shown here) produced about the same response in the Southern as in the Northern Hemisphere. The larger Southern Hemisphere thermal inertia almost exactly compensated for the larger forcing.

The effects of using the severe 10,000 MT forcing¹ are shown in run 8. The results during the first year are virtually the same as in the control case, with the maximum cooling not even 2 °C larger. The dramatic difference comes in the second year, where the snow/albedo feedback produces a cooling of more than 16 °C during the summer 1 yr after the war begins. The effects persist for several years after this, with year 4 in this case resembling year 2 in the control case. Although this result is almost a 'worst case' scenario (more latitudinal spread would make it even worse), it should receive serious consideration¹ because it is plausible. Run 9 shows that even the smaller 100 MT-city attack can have virtually the same effects as the control case, in agreement with Turco *et al.*¹

These experiments show that the climatic effects of a nuclear war might persist longer than previously calculated. The use of a model which includes snow and ice feedbacks and ocean response shows that the effects can be large even 1 yr after the war begins. Latitudinal spread of the nuclear smoke and dust can produce large cooling in the tropics where life is more sensitive to cooling due to the absence of a natural seasonal cycle⁷, but the cooling there is not as large as would be expected from a global average model.

The results presented here must be considered as preliminary, since the climate model used cannot consider many of the complex interactions that might result. These include the dynamical and radiative interactions between the atmospheric circulation and the smoke, the possibly patchy nature of the smoke distribution, long-wave radiative interactions, ocean circulation and mixed-layer depth responses, atmospheric hydrological cycle responses (including changes in cloudiness and changes in washout rates due to changes in precipitation), the effects of considering diurnal solar forcing (R. Cess, personal communication), and the effects of placing dust over the smoke rather than having them evenly mixed¹⁹. What the initial distribution of smoke and dust would be is also not well known and depends on untested assumptions about numbers and sizes of fires, number of particles produced per fire, initial rainout and washout, particle size distribution and coagulation rates, vertical distribution, targeting strategies and initial weather conditions. All these processes must be included to produce realistic results. In the absence of proof that these assumptions would drastically change the results, the prospect of a nuclear winter following virtually any scenario for a nuclear war must be taken very seriously.

This work has been supported by NSF grant ATM-8213184. I thank S. Schneider, P. M. Kelly and S. Warren for valuable comments, the NASA Goddard Laboratory for Atmospheric Sciences for computer time on the Amdahl V6, and Chuck Mulchi and Patricia Palasik for drawing the figure.

- MacCracken, M. C. *Pap. at 3rd Int. Conf. on Nuclear War*, Erice, (Lawrence Livermore National Laboratory, UCRL-89770, 1983).
- Aleksandrov, V. V. & Stenchikov, G. L. *The Proceeding on Applied Mathematics* (The Computing Centre of the USSR Academy of Sciences, 1983).
- Covey, C., Schneider, S. H. & Thompson, S. L. *Nature* **308**, 21–25 (1984).
- Thompson, S. L. *et al. Ambio* (in the press).
- Crutzen, P. J. & Birks, J. W. *Ambio* **11**, 114–125 (1982).
- Ehrlich, P. R. *et al. Science* **222**, 1293–1300 (1983).
- Robock, A. *J. Atmos. Sci.* **40**, 986–997 (1983).
- Sellers, W. D. *J. Appl. Met.* **12**, 241–254 (1973).
- Sellers, W. D. *J. Appl. Met.* **13**, 831–833 (1974).
- Robock, A. *Geophys. Int.* (in the press).
- Robock, A. *Science* **206**, 1402–1404 (1979).
- Robock, A. *Science* **212**, 1383–1384 (1981).
- Robock, A. *J. Volcan. geotherm. Res.* **11**, 67–80 (1981).
- Robock, A. *J. Atmos. Sci.* **35**, 1111–1122 (1978).
- Robock, A. *Pap. at XXVth COSPAR Meet.* (in the press).
- Warren, S. G. & Wiscombe, W. J. *Nature* (submitted).
- Manabe, S. & Stouffer, R. J. *J. geophys. Res.* **85**, 5529–5554 (1980).
- Thompson, S. L. *Climat. Change* **6**, 105–107 (1984).

Worldwide marine temperature fluctuations 1856–1981

C. K. Folland, D. E. Parker & F. E. Kates

Meteorological Office, London Road, Bracknell,
Berkshire RG12 2SZ, UK

Paltridge and Woodruff¹ found a warming of global mean sea-surface temperatures (SSTs) of ~1 °C in the first half of the present century, but there are doubts about its reality because spurious temperature fluctuations can arise from changes in measuring instruments^{2,3} and because inter-annual fluctuations may contain sufficient variance to allow the observed inter-decadal variations to fall within the limits of sampling error of a stationary series³. We present here the results of analyses of worldwide SSTs and night near-surface marine air temperature (MAT) for the period 1856–1981 with the aim of estimating the magnitudes of recent climatic fluctuations of temperature at the ocean surface, taking account of changes in observing procedures. Our results show a worldwide temperature fluctuation of range ~0.6 °C (in broad agreement with a preliminary analysis of global SST²), with the coldest period being centred around 1905–10 and the warmest occurring in the 1940s. The fluctuation has a similar magnitude to, and is nearly in phase with, climatic warmings and coolings near the surface of the Northern Hemisphere land masses for the period after 1900. Before 1900 the trends are sharply different.

The data (46 million non-duplicated SST data and 24 million non-duplicated night MAT data) were derived from the Meteorological Office Main Marine Data Bank⁴. Observations departing from the whole-period normal by more than ±6 °C for SST and ±10 °C for MAT were rejected. Relaxation of the limits for SST to ±9 °C had little effect even in areas of high natural variability. The observations were composited into

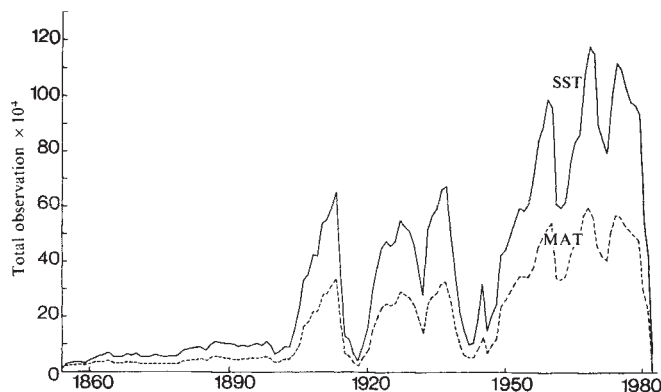


Fig. 1 Annual numbers of marine temperature observations for the globe. SST, sea-surface temperature; MAT, nighttime marine air temperature.

Received 2 April; accepted 26 June 1984.

I. Turco, R. P., Toon, O. B., Ackerman, T. P., Pollack, J. B. & Sagan, C. *Science* **222**, 1283–1292 (1983).

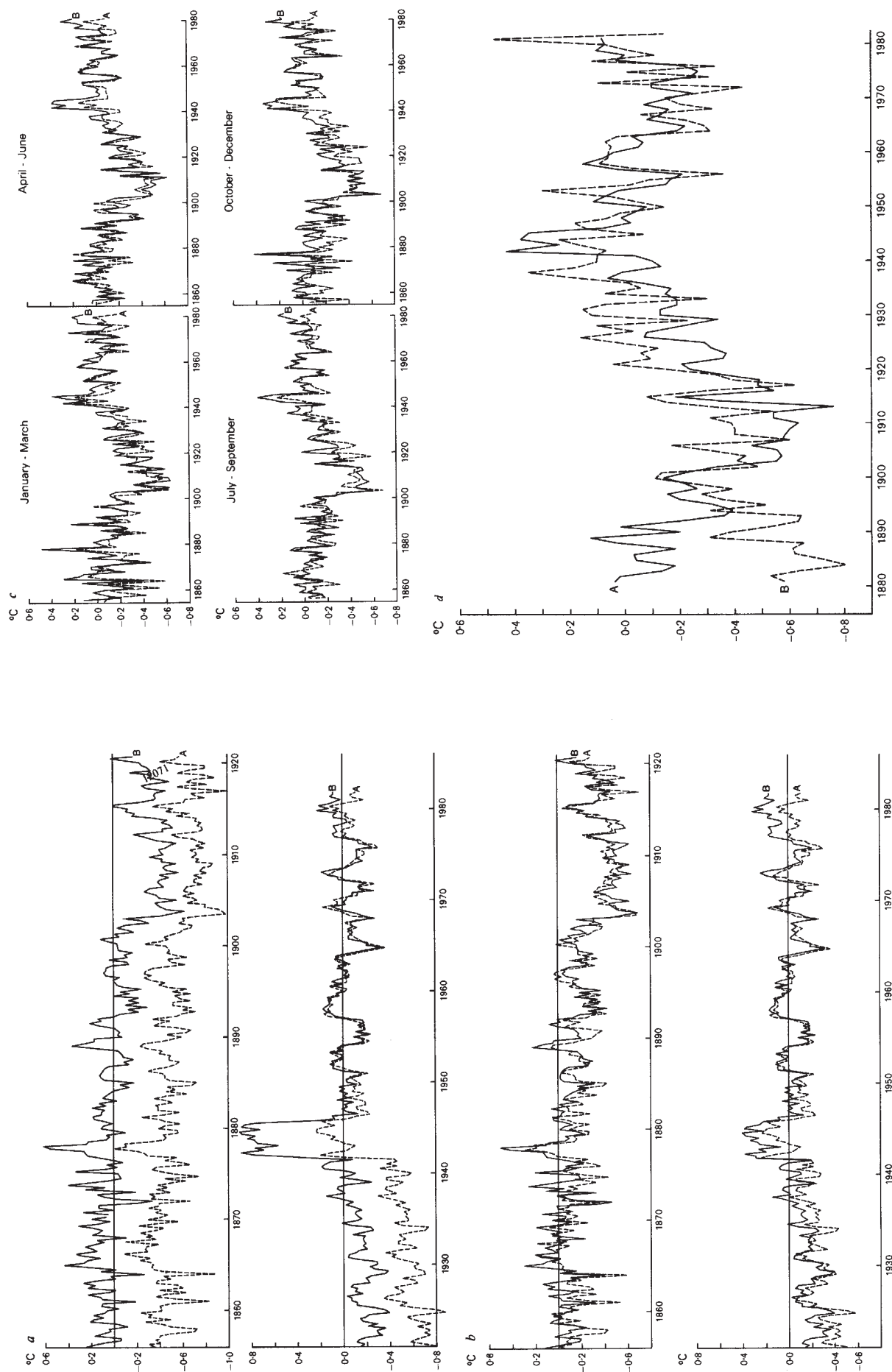


Fig. 2 *a*, Uncorrected anomalies (relative to 1951-60) of: A, global sea-surface temperature; and B, nighttime marine air temperature. *b*, Corrected anomalies (relative to 1951-60) of: A, global sea-surface temperature; and B, global nighttime marine air temperature. *c*, Corrected anomalies (relative to 1951-60) of: A, global sea-surface temperature; and B, global nighttime marine air temperature, for the seasons taken separately. *d*, Annual anomalies (relative to 1951-60) for the Northern Hemisphere of: A, corrected nighttime marine air temperature; and B, mainly land temperature after ref. 17.

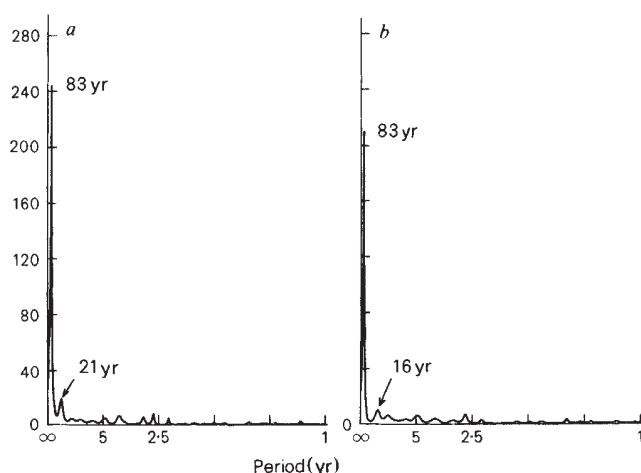


Fig. 3 Maximum entropy power spectra of corrected global seasonal marine air temperature (a) and sea-surface temperature (b), 1861–1980.

monthly anomalies from the 1951–60 normal for $5^\circ \times 5^\circ$ grid areas, taking account of the uneven spatial distribution of observations and also the space–time variations of the climatological average within each area. Areas with insufficient observations for these processes were omitted. Some areas of the Pacific with data missing in the 1960s were filled, after further quality control, with analysed data obtained from the Massachusetts Institute of Technology (MIT) (J. Hsiung, unpublished data). Note that the analysis of SST reported in ref. 2 used about 40 million data including some duplicates.

Daytime MAT observations on board ship are known to be affected by solar heating of the ship's fabric^{5–9}. Therefore, only MAT at night (solar elevation $<0^\circ$) was used. Local environmental effects cause MAT to be relatively too warm early in the night and slightly too cool later on², so that night MAT as a whole has little bias. The MIT data, being daily averages, could not be inserted into the night MAT data sets, but comparisons between SST and night MAT for the Pacific for the 1960s indicate that the gaps in night MAT will not prejudice conclusions on large-scale changes. Because of the gaps, the shorter reference period 1951–60, rather than a longer period, such as 1951–80, was used to calculate SST and MAT anomalies; the general characteristics of the time series of anomalies were found to be unaffected by this choice.

Figure 1 shows the annual numbers of SST and night MAT observations used. Note the reduction during the world wars. Graphs (not shown) for individual oceans are similar but there are more Northern Hemisphere than Southern Hemisphere data after 1880. Global coverage, measured as an areally weighted proportion of grid areas with monthly data, averaged 20% before 1900, nearly 60% between the world wars and over 80% in the 1970s. Monthly grid area anomalies were averaged over successive 3-month periods (January–March and so on) and the results averaged globally with appropriate areal weighting. These global 'seasonal' averages, uncorrected for changing instrumental and procedural biases, are presented sequentially in Fig. 2a. The SST and night MAT curves are not parallel and demonstrate the need to investigate instrumental and procedural changes.

The night MAT data are affected by the gradual increase in average elevation of thermometer screens above the sea from ~6 m before 1900 to >20 m in the most recent years (R. C. Cameron, personal communication) as ships increased in size. Corrections to reduce temperatures to the mean screen elevation (15 m) considered appropriate during 1951–60 were computed using atmospheric surface boundary layer similarity theory¹⁰, with an estimated global average Richardson number $Ri =$

-0.01 . This value was deduced from (1) monthly average fields of approximate air–sea-temperature differences, computed from night MAT and estimated night SST data with the aid of published data on the diurnal range of SST¹¹ and including a first-guess adjustment to MAT for changes in screen height; (2) a globally-averaged wind speed. The corrections were found to be insensitive to likely uncertainties in the value of Ri . The marine surface air layer is, on average, superadiabatic; the corrections to MAT reflect this and are: up to 1900, -0.13°C ; 1901–15, -0.07°C ; 1916–60, 0.0°C ; 1961–70, $+0.02^\circ\text{C}$; 1971–75, $+0.06^\circ\text{C}$; 1976–81, $+0.09^\circ\text{C}$.

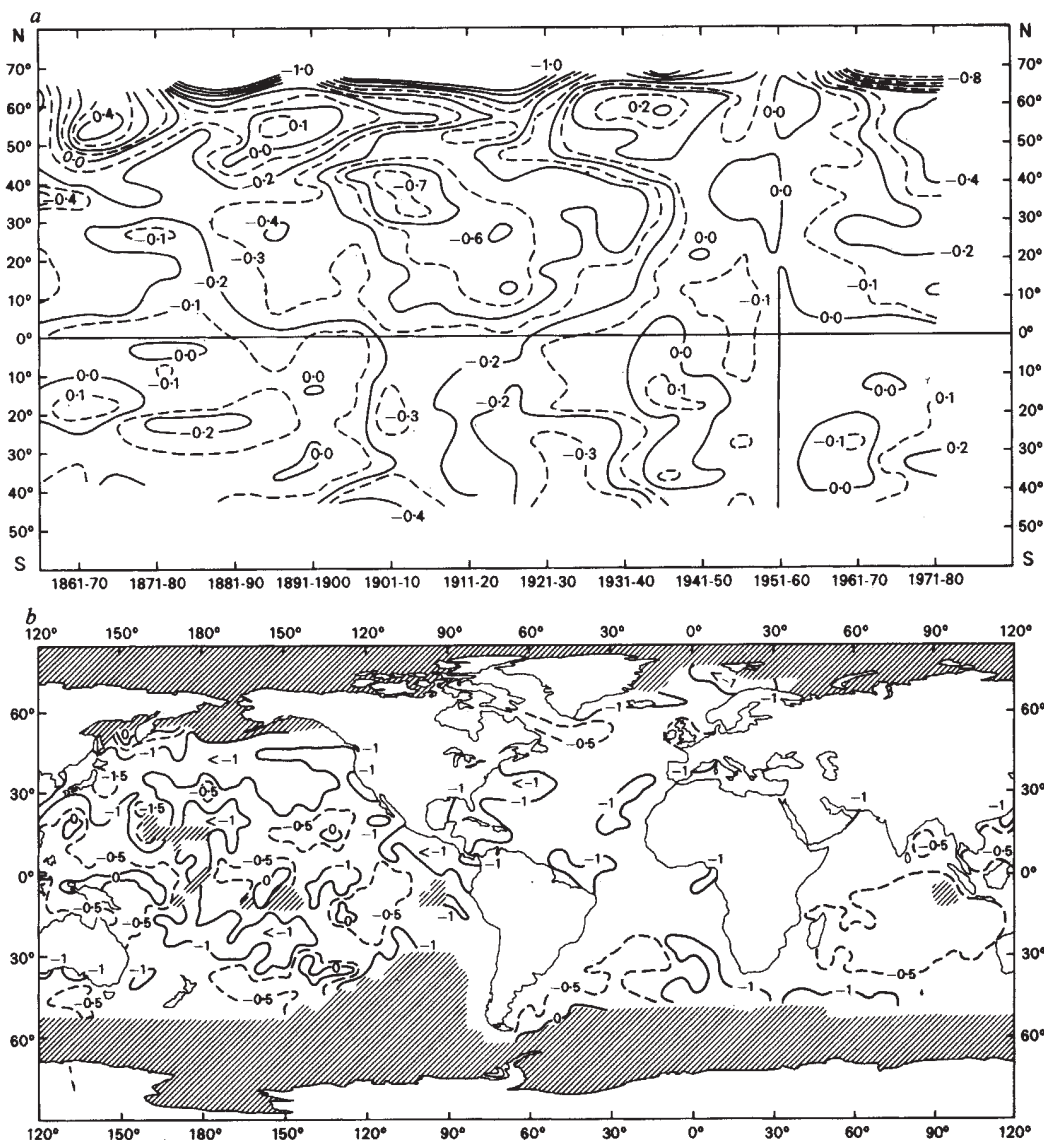
The extreme warmth in night MAT between 1942 and 1945 (Fig. 2a) merits investigation, especially as the SST series did not behave similarly. Comparison of sequences of day and night MAT during World War 2 showed that the warm episode was much less marked by day. The reason is thought to be that it was forbidden, at least on UK ships, to shine a torch in an exposed place, so night MAT was observed well in-board, with consequential larger heating errors (G. V. Mackie, personal communication). Daytime MAT was observed normally and was thus unaffected. The small peak in daytime MAT around 1942–45 may have been connected with the Southern Oscillation: there was a strong El Niño in 1940–41^{12,13}. Adjustments to night MAT were made assuming homogeneity of wartime day MAT, as follows: -0.1°C for April 1940 to the end of 1941; -0.5°C for January 1942 to September 1945, a period having more belligerents.

The change from uninsulated bucket measurements of SST to chiefly engine intake measurements seems to have taken place around the beginning of World War 2. Since World War 2, however, insulated buckets have been used extensively¹⁴, and in recent years globally-averaged SSTs derived from buckets and from engine intake probes appear from our own researches to differ by $<0.1^\circ\text{C}$. Early investigations using uninsulated buckets^{15,16} suggest that SST data derived from these instruments are likely to be depressed below their true magnitudes because of the cooling effect of evaporation from the wet walls. Thus the pre-World War 2 SST data require a positive correction to make them homogeneous with the 1951–60 reference period, whereas corrections to later SST data are likely to be small. We have used the corrected night MAT data as a reference and find that this yields a single positive correction of 0.3°C to the globally-averaged SST data before April 1940, a slightly smaller correction of 0.25°C between April 1940 and December 1941, and zero correction thereafter. This provides a consistent set of curves of global SST and night MAT (Fig. 2b). Note, however, that for regional studies the SST corrections needed may vary with season, because of annual cycles of dewpoint depression and wind speed which determine the evaporative effects.

The mid-nineteenth century values appear close to those of the 1951–60 reference period; there is a gradual fall of 0.2°C in the late nineteenth century followed by a sudden fall of about 0.3°C shortly after 1900; the warming of almost 0.6°C up to the 1940s is followed by a steady cooling of about 0.25°C in MAT, but the warmest decade in SST is 1951–60. Renewed warming appears to have set in after the early 1970s. The three-month 'seasons' show similar sequences (Fig. 2c). Although the time-lag of the long period variations of SST behind those of MAT is greatest ($>$ one decade) in the mid-twentieth century, it is not entirely absent in the earlier years. Except before 1900, the annual night MAT curve for the Northern Hemisphere is in good agreement with the annual Northern Hemisphere land curve presented by Jones *et al.*¹⁷ (Fig. 2d). The reality and physical significance of the apparent slight time-lag of the MAT curve relative to the land curve should be investigated.

The sequences in Fig. 2b were subjected to analysis of variance to test for the significance of inter-decadal changes when compared with within-decade 'seasonal' fluctuations. The F -ratio of 55.6 for night MAT was considered to have only 5 decadal and 50 within-decade degrees of freedom because of autocorrelation in the data¹⁸, but nonetheless has a probability of $<0.1\%$, so

Fig. 4 *a*, Zonally-averaged sea-surface temperature anomalies (relative to 1951–60) ($^{\circ}\text{C}$). Values are for 5-yearly overlapping decades and are corrected for instrumental changes. *b*, Corrected decadal sea-surface temperature for 1903–12 relative to 1951–60 ($^{\circ}\text{C}$). Hatching denotes missing data.



the null hypothesis of no inter-decadal climatic variation must be rejected confidently. Similar results hold for SST (F -ratio of 45.7 with 4 decadal and 44 within-decade degrees of freedom: probability $\ll 0.1\%$).

Detrended maximum entropy power spectral analyses on the sequences in Fig. 2*b* (Fig. 3*a*, *b*) both show a dominant peak at a period of 83 yr, expressing the major climatic fluctuation visible in Fig. 2*b*. Minor peaks at periods in the range 3–5 yr appear, according to work in progress, to be associated with the Southern Oscillation. An analysis (not shown) of Northern Hemisphere SST over the period 1945–80, when the long time scale climatic fluctuation was weak, shows fluctuations on a dominant time scale of 8–10 yr.

Figure 4*a* presents a latitude–time section of zonally-measured corrected SST from 65°N to 45°S . The major climatic fluctuation in the first half of the present century occurred simultaneously in both hemispheres but with greater amplitude in the Northern Hemisphere. The mid-latitudes of the two hemispheres now appear to be fluctuating out of phase, as they were to some extent before 1900. MAT gives quite similar results.

Figure 4*b* shows fields of difference in corrected SST between the globally warmest decade (1951–60) (anomaly 0.0°C) and the coldest (1903–12) (anomaly -0.44°C). The worldwide nature of the major fluctuations is clear. The corresponding warmest and coldest decades for night MAT are probably 1940–49

(anomaly 0.11°C) and more definitely 1903–12 (anomaly -0.50°C).

We consider that our results are the best estimates hitherto made of global climatic fluctuations of temperature at the ocean surface. They should provide valuable clues to the mechanisms of climatic change.

Received 13 March; accepted 6 June 1984.

1. Paltridge, G. & Woodruff, S. *Mon. Weath. Rev.* **109**, 2427–2434 (1981).
2. Folland, C. K. & Kates, F. E. (ed. Berger, A. L.) *Milankovitch and climate Part II*, 721–727 (1984).
3. Barnett, T. P. *Mon. Weath. Rev.* **112**, 303–312 (1984).
4. Shearman, R. J. *Met. Mag.* **112**, 1–10 (1983).
5. Folland, C. K. *Met. Mag.* **100**, 6–14 (1971).
6. Glahn, W. D. *Seewarte, Der Seewart* **2**, 250–256 (1933).
7. Goerss, J. S. & Duchon, C. E. *J. phys. Oceanogr.* **10**, 478–479 (1980).
8. Hayashi, S. *Jap. Met. Ag. J. Met. Res.* **26**, 84–87 (1974).
9. Walden, H. *Hamburg, D. Wetterd., Seewetteramt, Ann. Met.* **5**, 303–312 (1952).
10. Large, W. G. & Pond, S. *J. phys. Oceanogr.* **12**, 464–482 (1982).
11. Roll, H. U. *Int. Geophys. Ser.* **7** (Academic, New York, 1965).
12. Parker, D. E. *Met. Mag.* **112**, 184–188 (1983).
13. Pan, Y. H. & Oort, A. H. *Mon. Weath. Rev.* **111**, 1244–1258 (1983).
14. *International list of selected, supplementary and auxiliary ships No. 47* (World Meteorological Organization, 1955–1982).
15. Brooks, C. F. *Mon. Weath. Rev.* **54**, 241–254 (1926).
16. Lumby, J. R. *J. Cons. Perm. Int. Explor. Mer.* **3**, 340–350 (1928).
17. Jones, P. D., Wigley, T. M. L. & Kelly, P. M. *Mon. Weath. Rev.* **110**, 59–70 (1982).
18. Waldo-Lewis, R. P. & McIntosh, D. H. *Met. Mag.* **81**, 242–244 (1952).

Dynamics of meltwater discharge from Northern Hemisphere ice sheets during the last deglaciation

Richard H. Fillon & Douglas F. Williams

Department of Geology and Belle W. Baruch Institute, University of South Carolina, Columbia, South Carolina 29208, USA

During the last ice age, ice sheets covered large portions of North America, Europe, Asia, and the Arctic Ocean. The ice sheets were composed of meteoric water, enriched in the light isotope of oxygen (^{16}O), leaving oceanic seawater correspondingly enriched^{1,2} in ^{18}O . The volume of past continental ice sheets during the last glacial maximum 18,000 yr ago has been estimated from terrestrial evidence of ice limits³, glaciological modelling⁴, the $\delta^{18}\text{O}$ of foraminifera from deep sea cores^{5,6}, and sea-level evidence from raised coral terraces^{7,8}. Oxygen isotope studies of planktonic foraminifera reveal that $\delta^{18}\text{O}$ shifts during deglaciations in areas near sources of ice meltwater, like the Gulf of Mexico and Labrador Sea/Baffin Bay, are often larger than the corresponding average global oceanic shift⁹⁻¹¹ and that these variations reflect the timing of ice decay without the ocean-mixing lag contained in both distal and benthic $\delta^{18}\text{O}$ variations¹². We show here that geographically-distinct planktonic $\delta^{18}\text{O}$ gradients can be related to the volume of meltwater discharged through specific meltwater outlets and thus to the initial volume of the ice sheets at the commencement of deglaciation and to the subsequent dynamics of ice-sheet decay.

Use of $\delta^{18}\text{O}$ gradients provides volume estimates that do not require prior knowledge of the $\delta^{18}\text{O}$ of the meltwater and which are independent of salinity changes due to seasonal variations in sea-ice cover, because sea ice and its meltwater products are not isotopically depleted with respect to seawater^{13,14}. Neglecting sea-ice effects, lines of equal sea-surface salinity should approximately parallel seawater $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_{\text{sw}}$) contours because deglacial $\delta^{18}\text{O}_{\text{sw}}$ anomalies will primarily reflect mixing between seawater and an isotopically light freshwater dilutant. Actual salinity values, however, can be calculated only if the $\delta^{18}\text{O}$ of the dilutant is known or can be estimated independently.

For example, the correspondence between estimated and observed values of $\delta^{18}\text{O}_{\text{sw}}$ (Fig. 1) suggests that the $\delta^{18}\text{O}$ of planktonic foraminifera from dated cores, corrected for sea-surface palaeotemperature estimates¹⁵⁻¹⁸ (Table 1) can be used to reconstruct variations in $\delta^{18}\text{O}_{\text{sw}}$ for the last deglaciation. Errors of precision associated with deglacial $\delta^{18}\text{O}_{\text{sw}}$ estimates are about $\pm 0.6\%$, including analytical precision errors in the palaeotemperature estimates¹⁹ and the standard deviation involved in relating estimated $\delta^{18}\text{O}_{\text{sw}}$ to observed $\delta^{18}\text{O}_{\text{sw}}$ (Fig. 1).

High sedimentation rate cores from the Gulf of Mexico and Labrador Sea/Baffin Bay separately document the volume and timing of meltwater discharges from: (1) the southern margin of the Laurentide Ice Sheet through the Mississippi River and (2) the central, northern and eastern portions of the Laurentide Ice Sheet which must have drained principally through the Hudson Strait and Lancaster Sound during the later stages of the last deglaciation²⁰. Coverage of the meltwater discharges into the Arctic Ocean and adjacent seas from the northern fringes of the European, Asian and North American ice masses is provided by $\delta^{18}\text{O}$ records from the northern North Atlantic and Norwegian Sea.

Detailed stratigraphical studies of the cores employing ^{14}C dates, benthic foraminiferal $\delta^{18}\text{O}$ records and volcanic ash horizons^{15,21-27} strongly indicate that the peak $\delta^{18}\text{O}_{\text{sw}}$ anomalies (although not necessarily the peak discharge) coincide in both the Labrador Sea and Gulf of Mexico at ~ 13 kyr BP¹². This deep-sea stratigraphical evidence and terrestrial records of deglacial ice limits²⁰ together indicate that disintegration of the northern portion of the Laurentide Ice Sheet spanned the period

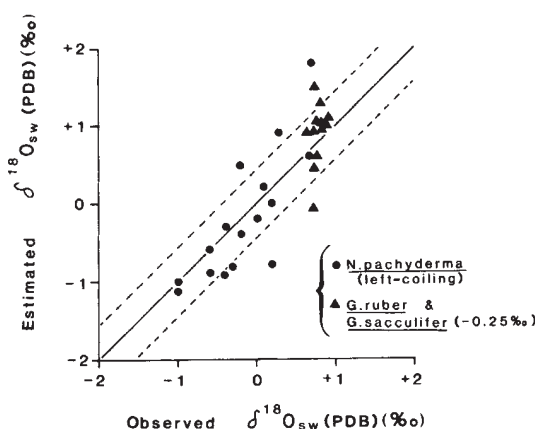


Fig. 1 Plot of seawater $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_{\text{sw}}$) estimates derived from the $\delta^{18}\text{O}$ of planktonic foraminiferal calcite²⁸ versus observed $\delta^{18}\text{O}_{\text{sw}}$ (ref. 43 and R. Fairbanks, personal communication) for North Atlantic and Gulf of Mexico core-tops (Table 1; Fig. 2). Dashed lines define a confidence band of $\pm$ one standard error of estimate.

~ 16 – 7 kyr BP and that meltwater discharge from the southern parts of the ice sheet into the Gulf of Mexico was limited to the period ~ 16 – 12 kyr BP¹².

In using planktonic foraminifera to assess meltwater-induced $\delta^{18}\text{O}_{\text{sw}}$ anomalies we assume that: (1) the oxygen isotopic temperature fractionation in the species selected for study, that is *Globigerinoides ruber*, *Globigerinoides sacculifer* and *Neogloboquadrina pachyderma* (left coiling), occurs in a fashion parallel to the equilibrium fractionation relationship between inorganic calcite and water ($\sim 0.22\%$ $\delta^{18}\text{O}/^{\circ}\text{C}$)^{28,29}; and (2) between-core variations in planktonic foraminiferal $\delta^{18}\text{O}$, corrected for between-site differences in observed sea-surface temperatures, rather than for temperatures at the depth of calcification of the foraminifera (0–100 m), yield $\delta^{18}\text{O}_{\text{sw}}$ estimates that are both close to and proportional to actual sea-surface $\delta^{18}\text{O}_{\text{sw}}$. Both assumptions have been tested (Fig. 1) by comparing $\delta^{18}\text{O}_{\text{sw}}$ estimates derived from core-top planktonic foraminifera with the observed $\delta^{18}\text{O}_{\text{sw}}$. In practice, in conditions of strong meltwater discharge, foraminiferal $\delta^{18}\text{O}$ anomalies probably yield minimum $\delta^{18}\text{O}_{\text{sw}}$ estimates because mixing of foraminiferal tests by bioturbation, strong carbonate dissolution and possible vertical movement of the foraminifera into higher salinities (isotopically heavier water) below a surface meltwater layer, would have diminished the apparent isotopic depletion as recorded in foraminiferal tests. With these constraints in mind, our estimates for sea-surface $\delta^{18}\text{O}_{\text{sw}}$ anomalies of 13 kyr BP are illustrated in Fig. 2a, b.

Records of benthic foraminiferal $\delta^{18}\text{O}$ indicate that global $\delta^{18}\text{O}_{\text{sw}}$ averaged about 1.2‰ heavier between 16 and 11 kyr BP relative to the modern seawater values^{22,27}. We, therefore, subtracted a value of 1.2‰ from planktonic foraminiferal $\delta^{18}\text{O}$ values in calculating the residual deglacial $\delta^{18}\text{O}_{\text{sw}}$ for the peak anomaly period 13 kyr BP shown in Table 1 and Fig. 2. Potential errors in estimating global $\delta^{18}\text{O}_{\text{sw}}$ enrichment at 13 kyr BP would be distributed equally among all the cores used in our study and so would not adversely affect $\delta^{18}\text{O}_{\text{sw}}$ gradients or the configuration and area of the deglacial $\delta^{18}\text{O}_{\text{sw}}$ anomaly. When so treated, $\delta^{18}\text{O}$ values for *G. ruber* in cores from the Caribbean Sea and tropical North Atlantic^{2,18}, at a considerable distance from meltwater sources, yield deglacial background $\delta^{18}\text{O}_{\text{sw}}$ estimates of $+0.5\%$ to $+0.3\%$, not too different from modern $\delta^{18}\text{O}_{\text{sw}}$ in those areas ($\sim 0.0\%$ to $+1.0\%$).

Comparisons of the areal extent of modern low salinity plumes (LSPs) due to freshwater runoff and the reconstructed 13 kyr BP approximate -0.6% $\delta^{18}\text{O}_{\text{sw}}$ contours (0‰ less a 0.6‰ precision error) show that the areal extent of the deglacial meltwater $\delta^{18}\text{O}$ anomaly was almost double the modern LSP in the Labrador

Table 1 Summary of core data and calculated $\delta^{18}\text{O}_{\text{sw}}$ estimates (Figs 1 and 2)

No.	Core	Water depth (m)	Core-top* $\delta^{18}\text{O}$ (‰)	Modern SST (°C)	13 kyr BP level*		Calculated $\delta^{18}\text{O}_{\text{sw}}$ (‰) vs PDB	
					$\delta^{18}\text{O}$ (‰)	SST estimates† (°C)	Core-top	13 kyr BP
1	HU76-40	2041	NA	-1.5 sp	+0.7‡ <i>N.p.</i> (26)	NA +2.0 sp (d)	NA	-3.9
2	HU77-17	935	NA	-1.5 sp	+0.8‡ <i>N.p.</i> (26)	NA +3.0 sp (d)	NA	-3.5
3	HU75-60	1,140	+2.2 <i>N.p.</i>	+2.9 sp	+3.8‡ <i>N.p.</i>	+6.0 sp (d)	-0.9	-0.1
4	HU75-58	1,057	+2.1 <i>N.p.</i>	+2.9 sp	+4.0‡ <i>N.p.</i>	+6.0 sp (d)	-1.0	+0.3
5	HU75-41	2,381	+2.7 <i>N.p.</i>	+1.0 sp	+3.7‡ <i>N.p.</i> (27)	+5.0 sp (d)	-0.8	-0.2
6	HU75-55	2,410	+2.1 <i>N.p.</i>	+3.5 sp	+1.4 <i>N.p.</i>	+5.0 sp (d)	-0.9	-2.5
7	HU75-31	2,505	+2.4 <i>N.p.</i>	+4.2 sp	+3.1 <i>N.p.</i>	+5.0 sp (d)	-0.4	-0.8
8	HU76-3	2,493	+2.0 <i>N.p.</i>	+3.0 sp	+2.0 <i>N.p.</i>	+6.0 sp (d)	-1.1	-1.7
9	HU75-4	3,731	+3.5 <i>N.p.</i>	+3.4 sp	+2.9 <i>N.p.</i>	+6.0 sp (d)	-0.4	-0.8
10	V28-14	1,855	+2.2 <i>N.p.</i>	+5.8 sp	+2.9 <i>N.p.</i> (21)	NA +3.0 sp (d)	-0.3	-1.4
11	V28-56	2,941	+2.9 <i>N.p.</i>	+3.3 sp	+3.3 <i>N.p.</i> (21)	NA +3.0 sp (d)	-0.2	-1.0
12	K-11	2,900	+3.0 <i>N.p.</i>	+3.6 sp	+3.2‡ <i>N.p.</i> (21)	NA +2.0 sp (d)	0.0	-1.3
13	KS7707	1,487	+3.7 <i>N.p.</i>	+2.8 sp	+3.9‡ <i>N.p.</i> (25)	NA +3.0 sp (d)	+0.5	-0.4
14	AII-94-PC4	2,114	+1.4 <i>N.p.</i>	+7.0 sp	+3.9‡ <i>N.p.</i>	+5.0 sp (c)	-0.8	+0.3
15	CH73-139C	2,209	+2.4 <i>N.p.</i>	+10.4 sp	+3.0‡ <i>N.p.</i> (22)	+9.0 sp (c)	+0.9	0.0
16	V27-116	3,202	+2.2 <i>N.p.</i>	+8.9 sp	3.3‡ <i>N.p.</i> (23)	+6.0 sp (c)	+0.2	-0.4
17	K708-1	4,053	NA	10.0 sp	+2.3 <i>N.p.</i> (23)	+8.0 sp (c)	NA	-0.9
18	V17-178	3,914	+1.7 <i>N.p.</i>	+6.7 sp	+2.1‡ <i>N.p.</i>	10.0 sp (c, e)	-0.6	-0.7
19	V30-101K	3,519	+2.5 <i>N.p.</i>	+14.0 sp	+2.0 <i>N.p.</i> (23)	+10.0 sp (c)	+1.8	-0.8
19	V30-101K	3,519	-0.2 <i>G.r.</i>	+20.6 su	+0.3‡ <i>G.r.</i> (23)	+16.0 su (b)	+0.6	-1.2
20	CH72-101	1,925	NA	+11.1 sp	+2.5‡ <i>N.p.</i> (22)	+10.0 sp (c)	NA	-0.3
21	CH72-104	1,755	NA	+11.3 sp	+2.5‡ <i>N.p.</i> (22)	+10.0 sp (c)	NA	0.0
22	CH67-19	1,800	NA	+10.5 sp	+2.5‡ <i>N.p.</i> (22)	+9.0 sp (c)	NA	-0.5
23	V27-86	2,900	+2.6 <i>N.p.</i>	+6.1 sp	+4.4‡ <i>N.p.</i> (24)	NA +3.0 sp (d)	+0.2	+0.1
24	NGI-31-36	2,620	NA	+7.8 sp	+3.5‡ <i>N.p.</i> (42)	NA +3.0 sp (d)	NA	-0.8
25	NGI-31-33	1,580	+1.6 <i>N.p.</i>	+7.8 sp	+2.7‡ <i>N.p.</i> (42)	NA +3.0 sp (d)	-0.4	-1.6
26	TAM12	1,600	-1.3 <i>G.r.</i>	+28.9 su	-1.5‡ <i>G.r.</i>	+27.0 su (a)	+1.3	-0.5
27	K97	3,408	-1.3 <i>G.s.</i>	+28.9 su	-2.6‡ <i>G.s.</i> (9)	+28.0 su (a)	+1.1	-1.6
28	K120	2,537	NA	+28.8 su	-2.3‡ <i>G.s.</i> (9)	+27.0 su (a)	NA	-1.6
29	TR126-29	2,700	-2.1 <i>G.r.</i>	-28.8 su	-2.9‡ <i>G.r.</i>	+27.0 su (a)	+0.5	-1.9
30	TR126-23	2,410	-1.5 <i>G.r.</i>	+28.8 su	-1.3‡ <i>G.r.</i>	+27.0 su (a)	+1.1	-0.4
31	K139	2,462	-1.4 <i>G.s.</i>	+28.8 su	-2.2‡ <i>G.s.</i> (9)	+27.0 su (a)	+1.0	-1.5
32	GS7603-10	890	-1.8 <i>G.r.</i>	+29.0 su	-0.5‡ <i>G.r.</i>	+28.0 su (a)	+1.5	+0.7
33	GS7603-13	1,600	-2.1 <i>G.r.</i>	+28.9 su	-3.2‡ <i>G.r.</i>	+28.0 su (a)	+0.5	-2.0
34	GS7102-9	695	-1.6 <i>G.r.</i>	+28.8 su	-2.2‡ <i>G.r.</i> (10)	+27.0 su (a)	+1.0	-1.3
35	GS7102-7	1,569	-1.7 <i>G.r.</i>	+29.0 su	-3.4‡ <i>G.r.</i> (11)	+27.0 su (a)	+0.9	-2.5
36	EN32PC-6	2,280	-1.6 <i>G.r.</i>	+29.1 su	-2.8‡ <i>G.r.</i> (12)	+27.0 su (a)	+1.0	-1.9

* Where possible averages of two or more samples have been taken to represent core-top (late Holocene) and 13 kyr BP $\delta^{18}\text{O}$. NA, Not analysed; *N.p.*, *Neogloboquadrina pachyderma* (left-coiling); *G.r.*, *Globigerinoides ruber*; *G.s.*, *Globigerinoides sacculifer*. (References are given in parentheses.)

† Sea surface temperature (SST) estimates for Spring (sp) and Summer (su) foraminiferal bloom periods have been determined as follows: (a) contoured trends from ref. 17; (b) ref. 41; (c) interpolated from refs 15, 16 using 13/10 kyr BP and 13/18 kyr BP SST relationships indicated in refs 22, 41; (d) contoured trends from unpublished SST estimates for cold subpolar waters extracted from discriminant and transfer functions computed by R.H.F. from core-top foraminiferal data¹⁶; (e) Winter SST=10°C, computed as in ref. 16 (W. Balsam, personal communication).

‡ Recognized as negative peak or step $\delta^{18}\text{O}$ anomaly.

§ No obvious $\delta^{18}\text{O}$ anomaly, 13 ± 1 kyr BP level interpolated as in refs 12, 15, 16, 23, 27.

Sea-Baffin Bay-North Atlantic region influenced by the deglacial Canadian and East Greenland current systems (Fig. 2a) and approximately five times the area of the modern LSP in the deglacial Gulf of Mexico (Fig. 2b). The area of the $\delta^{18}\text{O}_{\text{sw}}$ anomalies should be directly related to the area of the meltwater-induced low salinity plume regardless of the actual seawater salinity which, in the case of the North Atlantic, could have been lowered by low salinity water from melting sea ice¹⁴, and in the Gulf of Mexico, could have been enhanced by strong evaporation¹⁷.

Furthermore, an isotopically heavy dilutant, about -15‰³⁰ would imply very steep salinity gradients and minimum salinities, for example as low as ~24‰ ~6‰ less than modern values in Baffin Bay, where estimated $\delta^{18}\text{O}_{\text{sw}}$ at 13 kyr BP is as low as -4‰. On the other hand, an isotopically-light dilutant, about -33‰³¹ requires salinity gradients that are less steep and generally higher salinities, for example, of ~30‰ in Baffin Bay (≅ modern). By analogy with the modern LSPs, regardless of the actual salinities within deglacial LSPs, salinity/ $\delta^{18}\text{O}_{\text{sw}}$ gradients at the peripheries of the deglacial LSPs were probably sufficiently steep and well defined to cause reconstructed $\delta^{18}\text{O}_{\text{sw}}$

anomaly areas to be little affected by factors (such as bioturbation and minor changes in calcification depth) which would tend to decrease the magnitude of the meltwater $\delta^{18}\text{O}_{\text{sw}}$ anomaly.

A method has been devised to translate the area of the meltwater LSP inferred from $\delta^{18}\text{O}_{\text{sw}}$ contours into an equivalent rate of discharge of isotopically depleted freshwater. Using published hydrological and oceanographic data for 12 river discharge systems³²⁻³⁷, we found a strong linear correlation ($r = +0.98$) between rates of freshwater discharge and the areas of corresponding LSPs in the modern ocean. This relationship is described by the regression equation $Q = 3.00 \times A - 15.63$ where Q = freshwater discharge ($10^4 \text{ km}^3 \text{ kyr}^{-1}$) and A is LSP area (10^4 km^2) bounded within a band of steepened salinity gradients. Figure 3 compares observed discharges with discharges estimated using the regression equation. Reliable estimates of meltwater discharge ($\pm 0.5 \times 10^6 \text{ km}^3 \text{ kyr}^{-1}$, that is ± 2 standard errors of estimate) are thus possible as a function of the expanded area of the 13 kyr BP deglacial LSPs as indicated in Fig. 2a, b. The strong correlation between discharge and LSP area suggests that possible variations in the average thickness of LSPs are

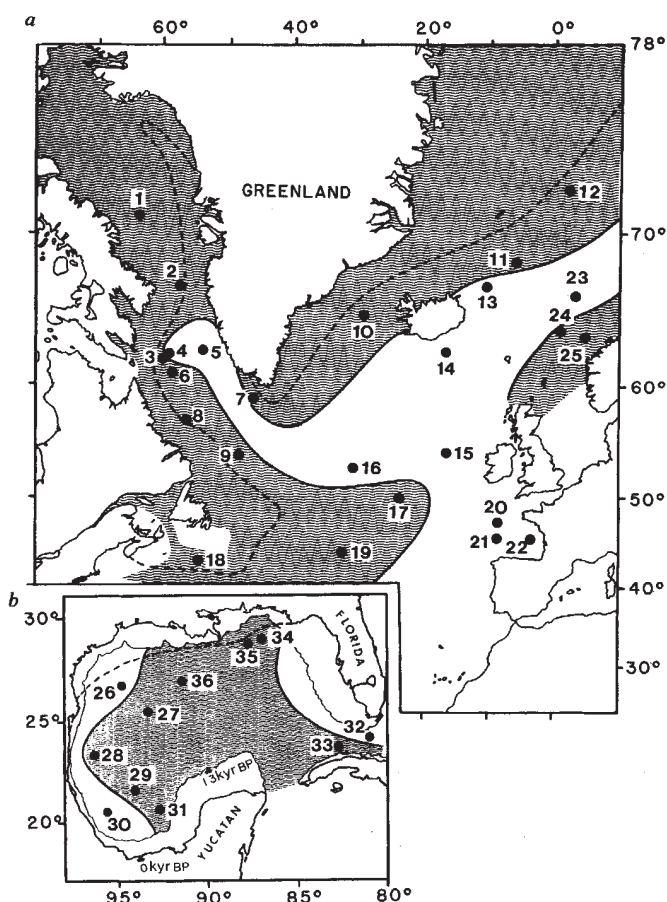


Fig. 2 Meltwater anomalies (shaded areas where estimated $\delta^{18}\text{O}_{\text{sw}}$ is $< -0.6\%$, that is, significantly $< 0\%$) in the northern North Atlantic (a) and Gulf of Mexico (b) ~ 13 kyr BP. Estimated $\delta^{18}\text{O}_{\text{sw}}$ variation is based on published and new $\delta^{18}\text{O}$ values of the planktonic foraminifera *Neoglobobulimina pachyderma* (left-coiling), *Globigerinoides ruber* and *Globigerinoides sacculifer* in 36 deep-sea cores (Table 1; refs 9–12, 21–27 and 42).

Table 2 Estimated meltwater discharge (10^6 km^3) into North American marginal seas and the North Atlantic during the last deglaciation

Meltwater source	Discharge 13–12 kyr BP*	Duration (kyr BP)	Cumulative discharge*
Gulf of Mexico	1.6	16–12	8.1
Canadian Current	3.6	14–7	29.8
East Greenland Current†	3.2	14–7	26.6
		Total discharge	64.5

* In units of 10^6 km^3 as excess over modern freshwater discharge.

† Includes the expanded Norwegian Coastal Current.

Table 3 Estimated minimum $\delta^{18}\text{O}$ of average ice sheets at the initiation of deglaciation (~ 16 kyr BP)

Glacial–interglacial $\delta^{18}\text{O}$ signal (‰)*	$\delta^{18}\text{O}$ of ice sheet
1.2	–25%
1.4	–29%
1.6	–33%
1.8	–38%

Based on possible glacial–interglacial $\delta^{18}\text{O}$ signals and estimated melt-water discharge (Table 2).

* Global oceanic $\delta^{18}\text{O}_{\text{sw}}$ change as recorded in deep-sea benthic foraminifera.

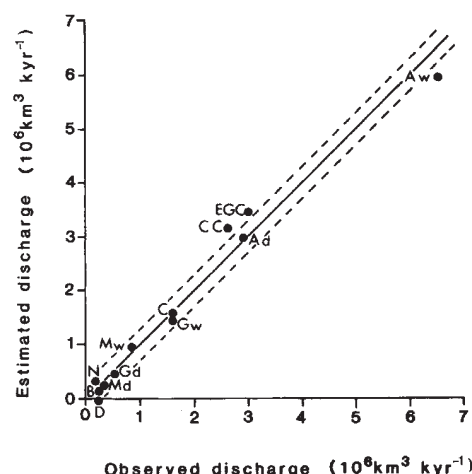


Fig. 3 Plot of freshwater discharge estimates based on the surface area of low salinity plumes in the modern ocean versus observed discharge values (refs 32–37) for 12 major freshwater discharge systems. Aw, Amazon River (wet season); Ad, Amazon River (dry season); EGC, East Greenland Current (mean); CC, Canadian Current (mean); C, Congo and Niger Rivers (mean); Gw, Ganges, Brahmaputra, Irrawaddy and Salween Rivers (wet season) and Gd (dry season); Mw, Mississippi River (wet season) and Md (dry season); N, Niger River (mean); B, combined Bosphorous outflow (mean); D, Danube, Dniestr and Dnepr Rivers (mean). Dashed lines define a confidence band of $\pm$ one standard error of estimate.

insignificant compared with the overall range of variation in the surface dimensions of LSPs ($\sim 10^4$ – 10^6 km^2 in Fig. 3).

Table 2 lists volume calculations for Laurentide meltwater discharged through the Canadian archipelago and Hudson Strait into the Canadian Current, for circum-Arctic Ocean meltwater discharged through the East Greenland Current, and for Laurentide meltwater discharged down the Mississippi River into the Gulf of Mexico during the interval from 13 to 12 kyr BP.

To estimate the total meltwater volume discharged during the entire deglacial interval (excess over modern freshwater discharge), modelled deviations of meltwater discharge (such as, Ruddiman and McIntyre's curve of estimated melt product influx²³) can be used with our 13 kyr BP estimates to extrapolate cumulative discharges for the period from 16 to 7 kyr BP (Table 2), which should thus approximate the volume of fresh water stored in Northern Hemisphere continental and marine ice sheets of the last glacial maximum. Incremental meltwater discharge (excess of runoff over accumulation) in this case is assumed to be inversely related to foraminiferal productivity in the North Atlantic²³.

If the Ruddiman and McIntyre model is basically accurate in its assumption that about 50% of Northern Hemisphere ice had been lost by 13 kyr BP³⁸ then for the period ~ 16 –12 kyr BP, it is estimated that the Gulf of Mexico received a total of $\sim 8.1 \times 10^6 \text{ km}^3$ of meltwater or $\sim 13\%$ of the total discharges. The subtotal discharge ($37.9 \times 10^6 \text{ km}^3$) into the Gulf of Mexico and Baffin Bay/Labrador Sea (Mississippi River plus Canadian Current) exceeds Flint's³ estimate for the volume of the Laurentide Ice Sheet by $\sim 10 \times 10^6 \text{ km}^3$, which is approximately the anomaly areas to be little affected by factors (such as bioturbation and minor changes in calcification depth) which would tend to decrease the magnitude of the meltwater $\delta^{18}\text{O}_{\text{sw}}$ anomaly.

The remaining meltwater ($\sim 26.6 \times 10^6 \text{ km}^3$) discharged into the Norwegian Sea and North Atlantic through the East Greenland Current apparently originated from European and Asian ice masses which drained into the Arctic Ocean. A portion of the meltwater attributed to the East Greenland Current (up to a maximum of $\sim 10 \times 10^6 \text{ km}^3$) may have been derived from an Arctic Ocean ice sheet^{38,39}. McKenzie River discharge from the

northwestern Laurentide ice sheet would also have contributed to some of the East Greenland Current discharge. A preliminary evaluation of Mediterranean $\delta^{18}\text{O}$ records points to only a minor contribution of glacial meltwater from rivers draining into the Black Sea, in agreement with the conclusions of other workers⁴⁰. The volume estimates contained in Table 2 should be considered minimum estimates, however, because potential contributions of meltwater from the Arctic Ocean through the Bering Strait (after its submergence) are not included in this analysis.

Simple calculations using the volume estimates in Table 2 place limits on the minimum $\delta^{18}\text{O}$ of the ice sheets (Table 3) for a probable range of values suggested as representing the glacial-interglacial change in $\delta^{18}\text{O}_{\text{sw}}$ (ref. 41). If significant meltwater contributions can be traced to the Bering Strait and to Antarctic deglaciation, the values in Table 3 would be increased proportionately (made less negative). The mean $\delta^{18}\text{O}$ of the ice sheets cannot be taken to represent the average $\delta^{18}\text{O}$ of the meltwater dilutant during a specific time interval, however, because of possible large variations in $\delta^{18}\text{O}$ within the ice sheets³⁸ and unknown variations in the $\delta^{18}\text{O}$ of precipitation during deglaciation. While we are therefore at present unable to accurately estimate salinity from estimated $\delta^{18}\text{O}_{\text{sw}}$, we feel the approach adopted here makes it possible to better understand the dynamics of meltwater discharge during the last, and possibly previous, deglaciations of the Pleistocene.

Received 14 March; accepted 6 June 1984.

- Emiliani, C. *J. Geol.* **74**, 109–126 (1966).
- van Donk, J. *Geol. Soc. Am. Mem.* **145**, 147–163 (1976).
- Flint, R. F. *Glacial Geology and the Pleistocene Epoch* (Wiley, New York, 1947).
- Hughes, T. J. et al. in *The Last Great Ice Sheets* (eds Denton, G. H. & Hughes, T. J.) 275–295. (Wiley, New York, 1981).
- Shackleton, N. J. *Nature* **215**, 15–17 (1967).
- Emiliani, C. *J. Geol.* **74**, 109–124 (1966).
- Chappell, J. *Bull. geol. Soc. Am.* **85**, 553–570 (1974).
- Fairbanks, R. G. & Matthews, R. K. *Quat. Res.* **10**, 181–196 (1978).
- Kennett, J. P. & Shackleton, N. J. *Science* **188**, 147–150 (1975).
- Emiliani, C. et al. *Science* **189**, 1083–1088 (1975).
- Emiliani, C., Rooth, C. & Stipp, J. J. *Earth planet. Sci. Lett.* **41**, 159–162 (1978).
- Leventer, A., Williams, D. F. & Kennett, J. P. *Earth planet. Sci. Lett.* **59**, 11–17 (1982).
- Craig, H. & Gordon, L. I. in *Stable Isotopes in Oceanographic Studies and Paleotemperatures, Spoleto 1965* (ed. Tongioli, E.) 9–130. (Consiglio Nazionale delle Ricerche, Laboratorio di Geologia Nucleare, Pisa, 1965).
- Tan, F. C. & Strain, P. M., *J. geophys. Res.* **85**, 1925–1932 (1980).
- Ruddiman, W. F. & Glover, L. K. *Quat. Res.* **5**, 361–389 (1975).
- McIntyre, A. et al. *Geol. Soc. Am. Mem.* **145**, 43–76 (1976).
- Brunner, C. A. thesis, Univ. Rhode Island (1978).
- Imbrie, J., van Donk, J. & Kipp, N. G. *Quat. Res.* **3**, 10–38 (1973).
- Molfini, B., Kipp, N. G. & Morley, J. J. *Quat. Res.* **17**, 279–313 (1982).
- Prest, V. K., in *Geology and Economic Minerals of Canada* (ed. Douglas, R. J. W.) 675–764 (Department Energy, Mines and Resources, Canada, 1970).
- Kellog, T. B., Duplessy, J. C. & Shackleton, N. J. *Boreas* **7**, 61–73 (1978).
- Duplessy, J. C., Delibrias, G., Turob, J. L., Pujol, C. & Duprat, J. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **35**, 121–144 (1981).
- Ruddiman, W. F. & McIntyre, A. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **35**, 145–214 (1981).
- Streeter, S. S., Belanger, P. E., Kellogg, T. B. & Duplessy, J. C. *Wuat. Res.* **18**, 72–90 (1982).
- Grousset, F. & Duplessy, J. C. *Mar. Geol.* **52**, M11–M17.
- Aksu, A. E. thesis, Dalhousie Univ. (1980).
- Fillon, R. H. & Duplessy, J. C. *Can. J. Earth Sci.* **17**, 831–854 (1980).
- Epstein, S., Buchsbaum, R., Lowenstam, H. & Urey, H. C. *Bull. geol. Soc. Am.* **64**, 1315–1325 (1953).
- Williams, D. F., Be, A. W. H. & Fairbanks, R. G. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **33**, 71–102 (1981).
- Emiliani, C. *J. Geol.* **63**, 538–578 (1955).
- Olausson, E., in *Progress in Oceanography* Vol. 3 (ed. M. Sears) 3, 221–252 (Pergamon, Oxford, 1965).
- Mosby, H. *Proc. Arctic Basin Symp.*, 69–83 (Arctic Institute of North America, 1963).
- van der Leeden, F. *Water Resources of the World: Selected Statistics* (Water Information Center, New York, 1975).
- Oceanographic Atlas of the Polar Seas*, Part 2 (H. O. Publ. 705, US Naval Oceanographic Office, Washington, 1958).
- Oceanographic Atlas of the North Atlantic Ocean, Section 2, Physical Properties* (H.O. Publ. 700, US Naval Oceanographic Office, Washington, 1967).
- Tchernia, P. in *Pergamon Marine Series* Vol. 3 (ed. Swallow, J. C.) (Pergamon, Oxford, 1980).
- Wyrtki, K. *Oceanographic Atlas of the International Indian Ocean Expedition* (National Science Foundation, Washington, 1971).
- Mix, A. C. & Ruddiman, W. F. *Quat. Res.* **21**, 1–20 (1984).
- Williams, D. F., Moore, W. S. & Fillon, R. H. *Earth planet. Sci. Lett.* **56**, 157–166 (1981).
- Rossignol-Strick, M., Nesteroff, W., Olive, P. & Bergnaud-Grazzini, C. *Nature* **295**, 105–110 (1982).
- Ruddiman, W. F. & McIntyre, A. *Science* **121**, 617–627 (1981).
- Sejrup, H. P., Eystein, J., Erlenkeuser, H. & Holteidahl, J. *Quat. Res.* **21**, 74–84 (1984).
- Tan, F. C. & Strain, P. M. *Bedford Inst. Oceanogr. Rep. B1-d-80-1*, (1980).

Sulphur-rich volcanic eruptions and stratospheric aerosols

Michael R. Rampino*[†] & Stephen Self[‡]

* NASA Goddard Institute for Space Studies, New York, New York 10025, USA

† Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

‡ Department of Geology, University of Texas at Arlington, Arlington, Texas 76019, USA

During the past decade it has become clear that the long-lived stratospheric clouds produced by volcanic eruptions are composed largely of sulphuric acid aerosols^{1,2}. The amount of sulphur-rich volatiles (for example, SO_2 , H_2S) injected into the stratosphere by an explosive eruption is, therefore, a critical determinant of its atmospheric impact^{3,4}. The small-volume eruptions of Mt Agung in 1963 and El Chichón in 1982 both generated substantial stratospheric aerosol clouds, despite the fact that they erupted $\leq 0.5 \text{ km}^3$ of magma. Comparison of data from direct measurements of stratospheric optical depth, Greenland ice-core acidity, and volcanological studies shows that such relatively small, but sulphur-rich, eruptions can have atmospheric effects equal to or even greater than much larger sulphur-poor eruptions. These small eruptions are probably the most frequent cause of increased stratospheric aerosols.

The recent eruption of El Chichón has provided fresh evidence that small-volume explosive events can have significant atmospheric impact^{5,6}. An older, but classic example is the 1963 eruption of Mt Agung (Gunung Agung) on the island of Bali, which is considered to be among the most important volcanic events of the twentieth century, primarily because of its possible effects on global climate^{7,8}. The 1963 Agung eruption (8°S) was a small-volume vulcanian to subplinian explosive event that produced about $0.3\text{--}0.6 \text{ km}^3$ of basaltic andesite magma. It has been given a volcanic explosivity index (VEI)^{9,10} rating of 4 on the scale of 0 to 7. The VEI is a combined measure of the explosive violence and volume of discharged pyroclastics of an eruption. The rather unusual Agung eruption sequence, with an early lava flow followed by climactic explosions, may be explained by the mixing of two distinct magma types¹¹. The duration of the major explosive phase on 17 March 1963 was about 7 h, while the 16 May eruption lasted for 5 h, suggesting peak rates of output of dense magma of $\sim 10^4 \text{ m}^3 \text{ s}^{-1}$. This is calculated to have produced eruption columns at least 18 km above sea level (a.s.l.)¹². The eyewitness reports of column heights on 17 March and 16 May were only 13 km a.s.l.¹³, but ash and sulphate aerosols were collected over Australia at 20 km a.s.l. from April 1963 up to about a year afterwards¹⁴. The resultant change in stratospheric aerosol optical depth¹⁵ peaked in the Southern Hemisphere subtropics and mid-latitudes at 0.2 to 0.3 from August to November 1963; Northern Hemisphere peak optical depths were much less, only ~ 0.02 measured at Hawaii¹⁶ during late 1963.

The measured optical-depth changes after the eruption suggest $\sim 1\text{--}2 \times 10^{13} \text{ g}$ of aerosol in the stratosphere^{17,18}. Based on the maximum estimated bulk volume of Agung ashfall ($\sim 1 \text{ km}^3$), the total amount of very fine ($< 2 \mu\text{m}$) ash that could have been produced by the eruption is roughly estimated¹⁹ as $< 8 \times 10^{12} \text{ g}$. Direct high-altitude sampling soon after the eruption suggests $\sim 5 \times 10^{12} \text{ g}$ of ash in the stratosphere²⁰. Such fine ash has only a brief stratospheric residence time, however, and was largely removed within a few months, leaving an aerosol cloud composed mainly of H_2SO_4 droplets²⁰.

For comparison, the El Chichón (17°N) eruption in 1982 was also small in volume, producing about $0.3\text{--}0.35 \text{ km}^3$ of trachyan-desite magma^{21,22}. The eruption had three main subplinian explosive phases of VEI = 4 on 28 March, 3 and 4 April. The eruption columns penetrated the stratosphere, where the main

Table 1 Comparison of eruption characteristics

Eruption	VEI ⁹	Volume of magma erupted (km ³)	Eruption column height (km) (c = calculated o = observed)	Stratospheric H ₂ SO ₄ aerosols (g) (from optical depth) ¹⁷	Stratospheric H ₂ SO ₄ aerosols (g) (from ice-core acidity) ²⁹	H ₂ SO ₄ yield (g) (from inclusion analysis) ²⁸	Northern Hemisphere ΔT (°C)	Refs (cols 3–8)
Tambora (8° S) 1815	7	≥50	>40 (c)	2×10^{14}	1.5×10^{14}	5×10^{13}	-0.4 to -0.7	35, 35, 17, 29, 28, 17
Krakatau (6° S) 1883	6	≥10	>40 (c)	5×10^{13}	5.5×10^{13}	3×10^{12}	-0.3	3, 3, 1, 29, 28, 36
Santa Maria (15° N) 1902†	6	~9	>30 (c, o)	$\leq 2 \times 10^{13}$	$\leq 2 \times 10^{13}$	—	-0.4	37, 37, †, †, —, 36
Katmai (58° N) 1912	6	15	>27 (c)	$\leq 2 \times 10^{13}$	$\leq 3 \times 10^{13}$	—	-0.2	38, 38, †, †, —, 36
Mt St Helens (46° N) 1980	5	0.35	22 (o)	$\sim 3 \times 10^{11}$	—	8×10^{10}	0 to -0.1	39, 40, 40, —, 28, 41
Agung (8° S) 1963	4	0.3 to 0.6	>18 (c) 13 (o)	$1-2 \times 10^{13}$	‡	3×10^{12}	-0.3	11, 11, 23, —, 28, 36
El Chichón (17° N) 1982	4*	0.3 to 0.35	26 (o)	$1-2 \times 10^{13}$	—	7×10^{10}	-0.4 to -0.6	22, 6, 23, —, 28, 42
Laki (64° N) 1783	4	0.3 (tephra) 12.3 (lava)	?	—	$\leq 1 \times 10^{14}$	9×10^{13}	~-1.0	4, —, —, †, 28, 43

* S. S., unpublished data.

† See text.

‡ Hammer *et al.*²⁹ give 2×10^{13} g, but this acidity peak cannot be a result of the Agung eruption, see text.|| Jones⁴² states that much of this decrease occurred between January and March 1982, before the eruption took place.

stratum of El Chichón aerosol spread out at about 26 km altitude⁶. Estimates based on lunar eclipse data²³ give an average global peak optical depth of 0.12 in December 1982; the aerosol was preferentially concentrated in the Northern Hemisphere, where the optical depth was 0.15. Conversion from optical depth to sulphuric aerosol mass loading¹⁷ yields about 2×10^{13} g of H₂SO₄ aerosol in the global stratosphere 6 months after the eruption. This agrees with estimates based on airborne lidar measurements²⁴ (1.2×10^{13} g) and on balloon-borne particle counters²⁵ ($1-2 \times 10^{13}$ g).

The large amount of sulphur volatiles released during these eruptions and injected into the stratosphere seem to be the most critical factor in their long-term atmospheric effects^{3,4}. Several studies have suggested that iron-rich basaltic and andesitic melts have a generally high capacity to carry dissolved sulphur^{4,26}. Melts of iron-poor dacite to rhyolite composition that commonly produce large volume, explosive plinian and ignimbrite eruptions (for example Krakatau 1883, Santa Maria, 1902) are usually poorer in sulphur volatiles^{3,27,28}.

Table 1 compares the vital statistics of significant explosive volcanic eruptions of the past 200 yr. The explosive magnitude of an eruption is succinctly summarized by its VEI value (column 2). The largest eruptions produced voluminous airflow and pyroclastic flow deposits (ignimbrite) (column 3) and had very high eruption columns (column 4). The global stratospheric aerosol mass loadings in column 5 were calculated using the method of Stothers¹⁷, where mass loadings (M_D) may be obtained from the global-average peak optical depth measurement (τ_D) using the formula $M_D = 1.5 \times 10^{14} \tau_D$ (grams). The estimates of global stratospheric aerosols based on Greenland ice-core acidities (column 6) are from Ref. 29. The values for Santa Maria, Katmai and Laki are given as upper limits because the aerosol clouds from these eruptions were probably restricted to the Northern Hemisphere or some portion of it. The sulphuric acid yield of the eruptions listed in column 7, taken from Ref. 28 are based on microprobe analyses of sulphur in volcanic ash samples and on estimates of the pre-eruption sulphur content of the magma from analyses of sulphur volatiles in glass inclusions trapped within crystals that formed just before the eruption²⁸. The values in column 7 were derived by subtracting the amount of sulphur remaining within the ash from the amount of sulphur volatiles found in the inclusions and then multiplying by the total volume of erupted magma. This 'sulphur yield' was then converted into the equivalent sulphuric acid aerosol mass²⁸.

The maximum observed post-eruption Northern Hemisphere temperature decreases that might be associated with the erup-

tions are listed in column 8. Specifically, ΔT is the difference between the temperature in the year before the eruption and the lowest temperature in the 1–3 yr following the eruption. In the case of St Helens, Northern Hemisphere temperatures rose in the 2 yr following the eruption, and the quoted decrease is the theoretical value of Robock⁴¹. However, not all of the temperature decreases can be attributed to the volcanic aerosols; in several instances the temperature decreases began before the eruption in question and in all cases, the signal is of the same order as interannual variability⁷. Furthermore, in certain cases, (for example, 1902) several eruptions capable of producing stratospheric aerosols may have occurred in close succession [for example, Pelée 1902 (VEI = 4) and Soufrière 1902 (VEI = 4)].

Several conclusions can be reached based on the data in Table 1. First, small magnitude volcanic eruptions can produce atmospheric after-effects like those of much larger eruptions. Second, there is good agreement between the aerosol estimates based on direct observations of changes in stratospheric optical depth^{3,17,18,23}, and those based on acid fallout over the Greenland ice sheet²⁹. In the case of Agung, however, the agreement must be fortuitous; stratospheric optical depth measurements clearly show that the bulk of the aerosol stayed in the Southern Hemisphere¹⁵. The small Northern Hemisphere optical depth of 0.02 indicates only $\sim 3 \times 10^{12}$ g of aerosol in the Northern Hemisphere stratosphere. The Greenland acidity peak in 1963–64 is too large, possibly because of input of aerosols from high northern latitude eruptions (such as, Surtsey in Iceland). Third, in all cases, the calculations of total sulphur release based on volcanological studies²⁸ are evidently underestimates, although in a relative sense they are in general agreement with the other methods. For El Chichón, the volcanological estimate of sulphur release is more than 2 orders of magnitude too low compared with the direct measurements of sulphuric acid aerosols in the stratosphere. Clearly these analyses are missing some source of additional sulphur release.

Where could the excess sulphur released in these eruptions have come from? In the case of El Chichón, anhydrite (CaSO₄; considered rare in volcanic rocks) occurs as a common phenocryst phase in the deposits, and decomposition of anhydrite has been suggested as a source of additional sulphur during the eruption^{22,28}. The Agung deposits, however, show no evidence of anhydrite. Another possible source of sulphur volatiles is degassing of non-erupted magma³⁰. Rose *et al.*^{30,31} calculated that for the 1974 eruption of Fuego (Guatemala) (also a small andesitic eruption) the amount of sulphur released required degassing of a body of intrusive magma about 5 times the mass

of the erupted material. If a similar situation occurred during the Agung eruption, then the total sulphur release would have been about 2.6×10^{12} g (or enough to produce about 8×10^{12} g of H_2SO_4 aerosols).

This study suggests that frequent, very-intense vulcanian to sub-plinian eruptions from the world's many active basaltic-andesitic composite volcanoes, such as Agung 1963, are probably the most common cause of significant perturbations of the stratospheric aerosol layer. These small eruptions are often rich in sulphur volatiles³¹ and can inject them efficiently into the stratospheric aerosol layer. The source of the sulphur enrichment, however, is not obvious. To solve this problem, a better understanding is required of the sulphur systematics in magmas and the processes taking place in shallow magma bodies beneath the volcanoes. For example, physical mixing of magmas of different composition may have been an important driving force in the Agung eruption and in many other basaltic-andesite volcanoes³²; such magma-mixing events may also control the concentrations and mode of release of sulphur gases and other volatiles during these eruptions.

A possible connection has been previously established between volcanic perturbations of the stratospheric aerosol layer and short-term (1–5 yr) climate cooling^{3,4,7,33,34}. We note from our new results in Table 1 that for the eruptions studied, ΔT is roughly proportional to $M_D^{1/3}$, as has been found previously by Devine *et al.*²⁸. Studies of polar ice cores show a further correlation between longer (decadal) periods of elevated ice-core acidity and cool climate²⁹. This suggests that episodes of increased sulphur-rich explosive volcanism may be an important modulator of long-term climate change. The connection between sulphur-rich volcanism and climate cooling may extend to large-volume basaltic fissure-type eruptions such as Laki (1783)^{4,28}. Large quantities of sulphuric acid aerosols could possibly be lofted into the lower stratosphere by such eruptions²⁸; ultimately, flood-basalt extrusions may have the greatest effects on the global atmosphere of any volcanic events.

We thank the Indonesian Volcanological Survey, Bandung, and especially Adjat Sudradjat (director) for hospitality and assistance, J. F. Luhr for unpublished analyses, H. Sigurdsson and R. B. Stothers for discussions, and R. S. J. Sparks and L. Wilson for critical reviews. The Indonesian Institute of Sciences (LIPI) granted permission to work on Bali in 1979. Supported by NASA grant NSG 5145 and cooperative agreement NCC 5–16 with Columbia University.

Received 3 January; accepted 15 June 1984.

- Castleman, A. W. Jr., Munkelwitz, H. R. & Manowitz, B. *Tellus* **26**, 222–234 (1974).
- Pollack, J. B. *et al.* *J. geophys. Res.* **81**, 1071–1083 (1976).
- Rampino, M. R. & Self, S. *Quat. Res.* **18**, 127–143 (1982).
- Sigurdsson, H. *EOS* **63**, 601–602 (1982).
- Pollack, J. B., Toon, O. B., Danielsen, E. F., Hofmann, D. J. & Rosen, J. M. *Geophys. Res. Lett.* **10**, 989–992 (1983).
- Rampino, M. R. & Self, S. *Scient. Am.* **205**, 48–57 (1984).
- Lamb, H. H. *Phil. Trans. R. Soc. A266*, 425–533 (1970).
- Hansen, J. E., Wang, W.-C. & Lacis, A. A. *Science* **199**, 1065–1068 (1978).
- Newhall, C. G. & Self, S. *J. geophys. Res.* **87**, 1231–1238 (1982).
- Simkin, T., Siebert, L., McClelland, L., Bridge, D., Newhall, C. & Latter, J. H. *Volcanoes of the World* (Hutchinson Ross, Stroudsburg, 1981).
- Self, S. & Rampino, M. R. Preprint, Univ. Texas (1984).
- Wilson, L., Sparks, R. S. J., Huang, T. C. & Watkins, N. D. *J. geophys. Res.* **83**, 1829–1836 (1978).
- Zen, M. T. & Hadikusumo, D. *Bull. volcan.* **27**, 269–299 (1965).
- Mossop, S. C. *Nature* **203**, 824–827 (1964).
- Volz, F. E. *J. geophys. Res.* **75**, 5185–5193 (1970).
- Mendonça, B. G., Hanson, K. J. & Deluisi, J. J. *Science* **202**, 512–515 (1978).
- Stothers, R. B. *Science* **224**, 1191–1198 (1984).
- Deirmendjian, D. *Adv. Geophys.* **16**, 267–297 (1973).
- Murrow, P. J., Rose, W. I. Jr & Self, S. *Geophys. Res. Lett.* **7**, 893–896 (1980).
- Toon, O. B. & Pollack, J. B. in *The Stratospheric Aerosol Layer* (ed. Whitten, R. C.) 121–147 (Springer, Berlin, 1982).
- Varekamp, J. C. & Luhr, J. F. *EOS* **63**, 1126 (1982).
- Duffield, W. A., Tilling, R. I. & Canul, R. J. *Volcan. Geotherm. Res.* **20**, 117–132 (1984).
- Keen, R. A. *Science* **222**, 1011–1013 (1983).
- McCormick, M. P. & Swissler, T. J. *Geophys. Res. Lett.* **10**, 877–880 (1983).
- Hofmann, D. J. & Rosen, J. M. *Univ. of Wyoming, Dept. Phys. Astr. Rep. AP-77*, 1–27 (1983).
- Haughton, D. R., Roeder, P. L. & Skinner, B. J. *Econ. Geol.* **69**, 451–467 (1974).
- Eichelberger, J. C. & Westrich, H. R. *Geophys. Res. Lett.* **8**, 757–760 (1981).
- Devine, J. D., Sigurdsson, H., Davis, A. N. & Self, S. *J. geophys. Res.* **89**, 6309–6325 (1984).
- Hammer, C. U., Clausen, H. B. & Dansgaard, W. *Nature* **288**, 230–235 (1980).
- Rose, W. I., Jr., Stoiber, R. E. & Maliniconco, L. L. in *Andesites* (ed. Thorpe, R. S.) 669–676 (Wiley, New York, 1982).

- Rose, W. I., Jr., Wunderman, R. L., Hoffman, M. F. & Gale, L. J. *Volcan. Geotherm. Res.* **17**, 133–157 (1983).
- Eichelberger, J. C. *Nature* **275**, 21–27 (1978).
- Taylor, B. L., Gal-Chen, T. & Schneider, S. H. *Q. J. R. met. Soc.* **106**, 175–199 (1980).
- Self, S., Rampino, M. R. & Barbera, J. J. *J. Volcan. Geotherm. Res.* **11**, 41–60 (1981).
- Self, S., Rampino, M. R., Newton, M. S. & Wolff, J. A. *Geology* (in the press).
- Jones, P. D., Wigley, T. M. L. & Kelly, P. M. *Mon. Weath. Rev.* **110**, 59–70 (1982).
- Williams, S. N. & Self, S. *J. Volcan. Geotherm. Res.* **16**, 33–56 (1983).
- Hildreth, W. J. *Volcan. Geotherm. Res.* **18**, 1–56 (1983).
- Lipman, P. W. & Mullineaux, D. R. (eds), *U.S. Geol. Survey Prof. Pap.* **1250**, 1–844 (1981).
- Newell, R. E. & Deepak, A. (eds) *NASA SP-458*, 1–119 (1982).
- Robock, A. *Science* **212**, 1383–1384 (1981).
- Jones, P. D. *Climate Monitor* **11**, 125 (1982).
- Köppen, W. Z. *Öst. Gesell. Met.* **8**, 257–267 (1973).

On the formation of calderas during ignimbrite eruptions

T. H. Druitt* & R. S. J. Sparks

Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK

Many large calderas result from the eruption of substantial volumes (tens or hundreds of km^3) of silicic pyroclastics. Such events often begin with an airfall phase and progress to the generation of voluminous ignimbrites^{1–3}. We propose here that many such eruptions involve two well-defined stages, based on a simple analysis of magma chamber pressure variations during an eruption. The first stage begins when an overpressured magma chamber fractures the country rock and forms a conduit to the surface. The chamber pressure decreases rapidly to values less than lithostatic pressure. We show that only small to moderate volumes of magma, representing a small fraction of the total chamber, can be erupted during this stage. In the second stage, caldera collapse results from a further decrease in magma pressure, which causes the chamber roof to fracture catastrophically and deform. Subsidence of the roof attempts to re-establish lithostatic pressures within the chamber and can drive substantial volumes of magma to the surface. Geological relationships in pyroclastic deposits associated with large caldera eruptions provide independent evidence for this model.

Present ideas on caldera formation are strongly influenced by studies in which it was recognized^{1,4} that collapse was fundamentally the consequence of removal of magma from the chamber. Smith and Bailey⁴ presented a model of caldera evolution, based largely on their studies of the Valles Caldera, New Mexico. They envisaged that an eruption began when a large body magma had accumulated within the upper crust, causing regional uplift and stretching. They also recognized that collapse often occurred while large volumes of ignimbrite were erupted from the chamber. The area of collapse is often proportional to the volume of erupted magma², suggesting a simple relationship between caldera collapse and removal of magma from the chamber.

The broad features of these classic ideas are widely accepted and are not in dispute here. We now present a simple analysis of magma chamber pressure during eruption, which suggests that caldera-forming events often have two distinct phases.

Figure 1 illustrates the two stages which are envisaged. Before eruption, magmatic overpressure becomes large enough to support the chamber roof and to propagate tensile or shear fractures to the surface^{4,5}, triggering an eruption. Smith and Bailey proposed that concentric, inward-dipping ring fractures were formed during the pre-eruption doming, and that these fractures provided pathways for initial magma ascent to the surface. Although there is evidence for early ring fracture development in some calderas^{6,7}, studies of the distribution of initial airfall deposits around volcanic centres^{1,8–13} show that eruptions can also begin in localized areas, perhaps along regional faults, which are not obviously connected to a ring fracture system. In

* Present address: US Geological Survey, MS-910, Menlo Park, California 94025, USA.

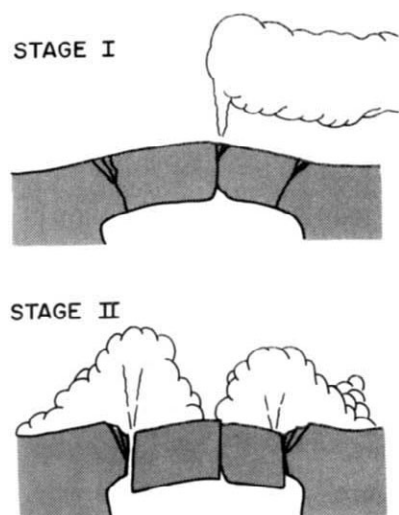


Fig. 1 Two stage model for caldera formation. In stage I, magma erupts from an initial vent while the chamber is overpressured. In stage II, the chamber pressure falls sufficiently for the chamber roof to collapse. New vents are thought to form at this stage. Other fracture geometries and collapse styles can be envisaged.

centres where there is a strong regional tectonic fabric, magma can preferentially exploit pre-existing structures during ascent^{8,14-16}, as well as early ring fractures. Walker¹⁶ has also shown that many post-caldera vents do not lie on ring-faults and that other lines of weakness, such as regional faults, often guide uprising magmas.

Once an eruption begins, removal of magma from the chamber will result in a decrease in pressure with time provided that the walls of the chamber remain rigid. Figure 2 shows a magma chamber with a volume, V_s , of volatile-rich silicic magma underlain by a volume, V_m , of volatile-poor magma. This configuration, in which the lower layer is generally more mafic, resembles the zoning envisaged in many natural magma chambers². The exact geometry of the chamber is not critical to the following arguments. In the following analysis, the silicic magma layer is assumed to be aphyric and to be just saturated with volatiles immediately before eruption at an overpressure of 25 MPa (250 bar), an upper limit for the tensile strength of a volcano⁵. We appreciate that the petrological evidence for pre-eruption saturation is equivocal and it is unlikely that the total volume of the chamber would be saturated. The emphasis here is on zoned chambers, but the treatment is a general one, which could also be applied to unzoned systems.

For a pressure drop, ΔP , the volume change ΔV of a magma with volume V is given by:

$$\frac{\Delta V}{V} = \frac{\Delta P}{\beta} \quad (1)$$

where β is the bulk modulus of the melt. The volume fraction (f) of the silicic magma layer (V_s in Fig. 2) discharged during an eruption can be related to the pressure change by the following expression:

$$f = x + \Delta P \left[\frac{1}{\beta_s} + \left(\frac{V_m}{V_s} \right) \left(\frac{1}{\beta_m} \right) (1 - x) \right] \quad (2)$$

where β_s and β_m are the bulk moduli of silicic and mafic magmas respectively; x is the volume fraction of vapour bubbles exsolved within the silicic magma layer and is given by:

$$x = \left[1 + \left(\frac{1 - N + S\sqrt{P}}{N - S\sqrt{P}} \right) \left(\frac{MP}{\sigma R \theta} \right) \right]^{-1} \quad (3)$$

where N is the weight fraction of volatiles in the magma, S the solubility constant, M the molecular weight of the volatile species, P the pressure, σ the melt density, R the molar gas

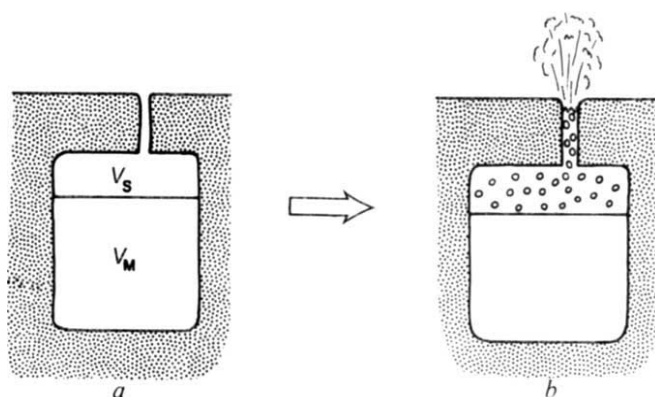


Fig. 2 The eruption of magma from a chamber of fixed volume containing two layers; an upper layer of water-saturated silicic magma with volume V_s and a lower layer, volume V_m , of water-poor magma.

constant and θ , the temperature. In the following calculations, the volatile species is assumed to be entirely water. The value of S is taken as 4.1×10^{-6} ($\text{Pa}^{-1/2}$) after Shaw¹⁷.

Solutions to equation (2) are plotted in Fig. 3, which shows the volume fraction of the silicic magma layer erupted as a function of magma pressure. The calculations assume that the chamber volume remains constant. Pressure is expressed in terms of departure from the lithostatic pressure at the roof of the magma chamber. The effect of considering pressure variations with depth in the chamber are small and are neglected. Curves are shown for magma chambers residing at depths of 2.5, 5 and 10 km in the crust. For each of these depths the lithostatic pressure has been calculated assuming a crustal density of $2,700 \text{ kg m}^{-3}$. The silicic magma layer at the top of the chamber is assumed to have an initial excess pressure of 25 MPa above the lithostatic value and to be just saturated in water at a temperature of 900°C . Using the solubility data of Shaw¹⁷ the implied water contents are 4.5, 5.7 and 7.5 wt % for 2.5, 5 and 10 km respectively. These values are perhaps realized only in the uppermost layers of chambers. Two curves are plotted for each depth in Fig. 3 corresponding to values of $V_m/V_s = 10$ and $V_m/V_s = 100$ (Fig. 2). This choice of values is based on geological² and theoretical¹⁸ deductions on the volumes of the volatile-rich caps to large silicic magma chambers. For each value of V_m/V_s curves are given for two values of β_m , the bulk modulus of mafic magma, of 10 GPa and 100 GPa. The value of β_s is assumed to be 30 GPa (ref. 19).

The major conclusion of these calculations is that only a small fraction (typically <0.3) of the silicic magma layer can be erupted from a magma chamber of fixed volume while an excess pressure is maintained (Fig. 3). For a given pressure drop the erupted fraction of the silicic magma layer increases as the chamber depth decreases. Consider a cylindrical magma chamber 11 km thick and 20 km in diameter situated 5 km deep in the crust. For $V_m/V_s = 10$, 6% of the silicic magma layer would discharge, 18 km^3 of magma, by the time the pressure had fallen to lithostatic. The solutions (Fig. 3) are also upper limits since we have assumed both initial water saturation of the magma and large initial overpressures. We conclude that in most silicic systems only small to moderate volumes, typically 0.1 km^3 to a few tens of km^3 in the largest systems, can be erupted while the chamber remains overpressured.

Magma can continue to be discharged from the chamber after the internal pressure decreases below lithostatic. A requirement for this to occur is that the eruptive conduit has a suitable geometry or is sufficiently irregular in shape to remain open. The main driving force for magma ascent is caused by the imbalance between the hydrostatic pressure of the vesiculating magma column within the conduit and the magma chamber pressure. The low density of the vesiculating magma results in the hydrostatic pressure being substantially less than the

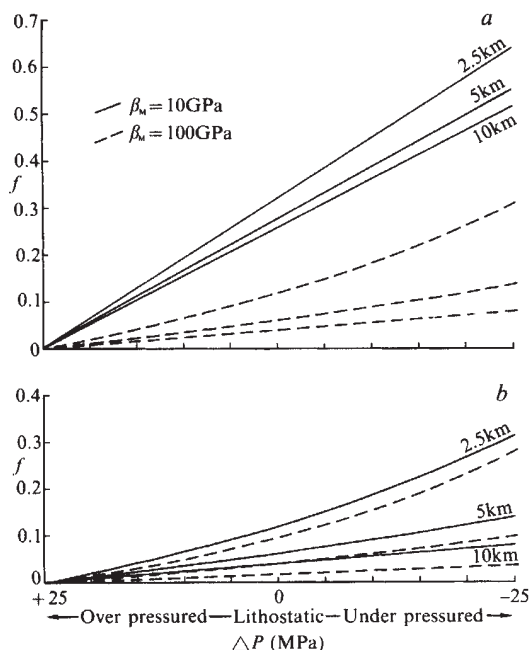


Fig. 3 The graphs show the volume fraction, f , of the silicic magma layer (V_s in Fig. 2) erupted as a function of decreasing chamber pressure. Departure of magma pressure from lithostatic pressure is illustrated assuming that the initial overpressure equals 25 MPa (250 bar). Solutions are shown for three chamber depths for $V_m/V_s = 100$ (a) and $V_m/V_s = 10$ (b). Solid line, $\beta_m = 10$ GPa; dashed line, $\beta_m = 100$ GPa.

chamber pressure even when the chamber pressure has fallen below the lithostatic value. The underpressure might be expected to become sufficient for the walls and roof of the chamber to collapse.

Initiation of the second stage of the eruption is marked by the onset of caldera collapse. Failure occurs because the chamber pressure becomes less than lithostatic, thereby relieving the roof of overlying support. Different styles of roof collapse can be envisaged. In one style, a coherent cauldron block bounded by outwardly-inclined ring fractures sinks into the chamber. Magma ascends along the fractures to form a ring dyke and erupts at the surface. In another style, the roof sags into the chamber along inwardly-inclined ring fractures or regional normal faults, perhaps deforming internally as collapse proceeds^{16,20}. In this case, magma need not necessarily ascend along the ring fractures, which will tend to close with collapse, but could either continue to erupt along the original conduit or open up new conduits as the roof fails. An interesting, but probably rare, variation could occur if collapse proceeds without any discharge of magma. In this case, subsidence will tend to restore the chamber pressure to lithostatic without further eruption. Regardless of collapse style, failure and subsidence of the roof rocks will attempt to re-establish lithostatic pressure in the chamber. Provided the conduits are not choked, hundreds or even thousands of km³ of magma could erupt during this stage.

Recent research on some pyroclastic deposits associated with caldera collapse supports the two-stage model. The deposit of the climactic eruption of Crater Lake, Oregon²¹ and the 18,500 yr BP Cape Riva eruption of Santorini^{10,22} show abrupt changes in the compositions and distribution of lithics and, in the Cape Riva case, of pumice clasts at the onset of major pyroclastic flow eruption. In both examples, an initial Plinian phase and minor pyroclastic flow activity occurred from a local vent. Subsequently the eruptions became much more violent and coarse co-ignimbrite lag breccias were discharged from new vent systems^{21,22}. G. P. L. Walker (personal communication) has also noted that proximal coarse lag breccias often rest on early pyroclastic flow deposits as well as the initial plinian deposit in many eruption sequences. The sudden appearance

of coarse lag breccias can be interpreted as marking the onset of caldera collapse, with the initial Plinian and pyroclastic flows representing the first stage of rapid chamber decompression.

Intra-caldera ignimbrite fills can often accumulate to thicknesses of 1–3 km within caldera depressions as subsidence occurs^{12–15,23}, providing perhaps the best evidence for synchronous caldera collapse during large magnitude eruptions. The volumes of ignimbrite trapped within caldera depressions can be hundreds of km³, often representing a large proportion of the ejecta.

We conclude that many caldera-forming eruptions initiate from chambers with excess pressure. However, only small to moderate volumes of magma can escape while an excess pressure is maintained. The products of the first stage include the initial Plinian deposit and perhaps some early ignimbrite units. Collapse occurs when the chamber pressure has decreased to substantially less than lithostatic and may often involve the formation of new vents. The abrupt appearance of coarse lag breccias in some eruption sequences may mark this transition to the second stage. Large volumes of ignimbrite can be erupted in this second stage, which sometimes infills the caldera depression as it forms.

We acknowledge useful comments from C. R. Bacon, S. Blake, R. Christiansen, P. Delaney, P. W. Lipman, R. MacDonald, H. Pinkerton, S. Self, G. P. L. Walker, L. Wilson and J. Wolff. R.S.J.S. was supported by the BP Venture Research Unit and T.H.D. by an NERC studentship, and subsequently by a NATO fellowship.

Received 13 March; accepted 8 June 1984.

- Williams, H. *Univ. Calif. Dept. Geol. Sci. Bull.* **25**, 239–346 (1941).
- Smith, R. L. *Geol. Soc. Am. Spec. Pap.* **180**, 5–27 (1979).
- Wilson, L., Sparks, R. S. J. & Walker, G. P. L. *Geophys. J. R. astr. Soc.* **63**, 117–148 (1980).
- Smith, R. L. & Bailey, R. A. *Geol. Soc. Am. Mem.* **116**, 613–662 (1968).
- Roberts, J. L. *Geol. J. Spec. Iss.* **2**, 287–338 (1970).
- Christiansen, R. L., Lipman, P. W., Orkild, P. P. & Byers, F. M. Jr. *US Geol. Surv. Prof. Pap.* **545B**, B43–B48, 1965.
- Christiansen, R. L. in *Explosive Volcanism, Inception, Evolution and Hazards*, 84–95 (National Academy Press, Washington DC, 1984).
- Simkin, T. & Fiske, R. S. *Krakatau 1883. The Volcanic Eruption and its Effects* (Smithsonian Institution Press, Washington DC, 1983).
- Sparks, R. S. J. *et al. J. Volcan. Geotherm. Res.* (in the press).
- Druitt, T. H. thesis, Univ. Cambridge (1983).
- Bond, A. & Sparks, R. S. J. *J. geol. Soc. Lond.* **132**, 1–16 (1976).
- Self, S. & Wright, J. V. *EOS* **62**, 1085 (1981).
- Walker, G. P. L. *J. Volcan. Geotherm. Res.* **8**, 69–94 (1980).
- Lipman, P. W. *J. geophys. Res.* (in the press).
- Francis, P. W. *et al. Nature* **301**, 51–53 (1983).
- Walker, G. P. L. *J. geophys. Res.* (in the press).
- Shaw, H. R. in *Geochemical Transport and Kinetics* (eds Hoffman, A. W., Giletti, B. J., Yoder, H. S. & Yund, R. A.) 139–170 (Carnegie Institute, Washington Publ. 634, 1974).
- Blake, S. *Nature* **289**, 783–785 (1981).
- Murase, T. & McBirney, A. R. *Bull. geol. Soc. Am.* **84**, 3563–3592 (1973).
- Reynolds, D. L. *Koninkl. Nederlandsch. Mijnb. Genootscap. Verh. Geol.* **16**, 355–398 (1956).
- Bacon, C. R. *J. Volcan. Geotherm. Res.* **18**, 57–115 (1983).
- Druitt, T. H. & Sparks, R. S. J. *J. Volcan. Geotherm. Res.* **13**, 147–171 (1982).
- Lipman, P. W. *Bull. geol. Soc. Am.* **87**, 1397–1410 (1976).

Selective population inversion in NMR

M. S. Silver*, R. I. Joseph*, C.-N. Chen†, V. J. Sank† & D. I. Hoult†

* Department of Electrical Engineering and Computer Science, The Johns Hopkins University, Baltimore, Maryland 21218, USA
† Biomedical Engineering and Instrumentation Branch, Division of Research Services, National Institutes of Health, Bethesda, Maryland 20205, USA

Population inversion of a selected region of a spectrum is a concept which has wide application in both NMR spectroscopy and imaging. While inversion of population at any one frequency is a trivial matter, ensuring an accurate inversion over a specified bandwidth, with negligible perturbation of the magnetization outside that bandwidth, is a major problem. However, by using as a driving function a complex radiofrequency (r.f.) pulse with an envelope

of the form $(\text{sech } \beta t)^{1+\mu}$ where $1/\beta$ is the temporal width and t is time, we have found that above a critical r.f. power threshold, magnetization is accurately inverted over a very sharply defined bandwidth, while outside that region, magnetization is returned to its initial position, and population is unaffected. Within the broad limits imposed by our equipment, we have also discovered that the phenomenon is independent of the incident r.f. power.

In general, the difficulty in producing a specific magnetization profile lies in the mathematical analysis of the problem—the differential equations (due to Bloch) which govern the motion of the magnetization in the presence of an irradiating magnetic field are nonlinear, rendering their inverse solution hard. In our search for a highly 'selective π pulse', we have considered the phenomenon of self-induced transparency in an analogous field of research—coherent optics. McCall and Hahn¹ have shown that above a critical power threshold, a pulse of light having an electric field described by a hyperbolic secant (h.s.), is transmitted through a normally absorptive medium with anomalously low energy loss. They have ascribed this phenomenon to the fact that a h.s. waveform produces a 2π rotation of polarization independent of the pulse power. Following the optical analysis of Lamb², one may apply analogous mathematics to the case of magnetic resonance. A full treatment is outside the scope of this letter. However, we indicate here the form of the analysis. An initial simplifying assumption is that the length of the pulse is very much less than the transverse relaxation time T_2 . The Bloch equations may then be written in the form

$$\begin{aligned} \dot{M} + i\Delta\omega M + iM_z\Omega &\approx 0 \\ \dot{M}_z + i(\Omega^* M - \Omega M^*)/2 &\approx 0 \end{aligned} \quad (1)$$

where M is the complex transverse magnetization $M_x + iM_y$, M_z is longitudinal magnetization, ω is the frequency and Ω is the driving function—the irradiating field $-\gamma(B_{1x} + iB_{1y})$. With the aid of the substitution

$$f = M/(M_0 + M_z) \quad (2)$$

the Bloch equations are reduced to a single Riccati equation

$$\dot{f} + i\Delta\omega f - i\Omega^* f^2/2 + i\Omega/2 \approx 0 \quad (3)$$

which is of considerable value for analysis with small and intermediate values of f , for f increases monotonically with increasing declination (flip angle) of the magnetization. However, for inversion, f becomes infinite. When the driving function is of the form

$$\Omega = \Omega_0(\text{sech } \beta t)^{1+\mu} \quad (4)$$

where β , Ω_0 and μ are real constants, use of the substitutions

$$2x = 1 + \tanh \beta t; \quad y = \exp\left\{-\frac{i}{2} \int_0^t f\Omega^* dt\right\} \quad (5)$$

enables one to reduce the Bloch–Riccati equation to the hypergeometric equation.

This reduction is performed by comparing the form

$$P(x) d^2y/dx^2 + Q(x) dy/dx + R(x)y = 0 \quad (6)$$

with equation (3). The hypergeometric series obtained as a solution is best related to observed magnetization by using the identity

$$M_z = M_0 \left\{ \frac{yy^*(\Omega_0/2\beta)^2 - x(1-x) dy/dx d^*y/dx}{yy^*(\Omega_0/2\beta)^2 + x(1-x) dy/dx d^*y/dx} \right\} \quad (7)$$

From equation (5), we may observe that the boundary conditions are established before the pulse at $x=0$, and the resultant is obtained at the end of the pulse as $x \rightarrow 1$. As the parameters Ω_0 , the amplitude of the pulse, $1/\beta$ its temporal width, and μ , its oscillatory nature, are varied, it is found that a highly selective π pulse is produced when $\mu \approx 5$ and $\Omega_0 \gg \mu\beta$. While the convergence of y in the range $0 \leq x < 1$ is absolute, that of dy/dx is conditional, and we are therefore investigating its behaviour in the belief that the high selectivity of the result is related to the stability of dy/dx as the distance from resonance $\Delta\omega$, is

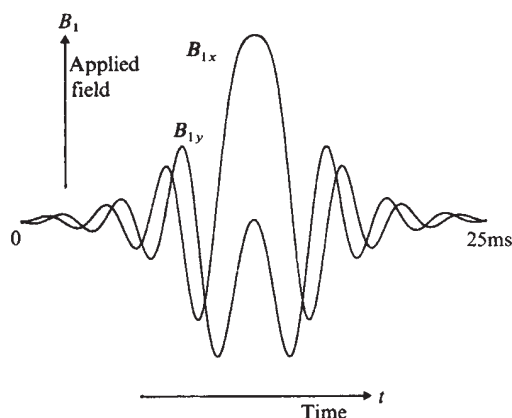


Fig. 1 The two components of the modulation used to create the complex hyperbolic secant pulse. In the rotating frame, the components are the irradiating x and y fields.

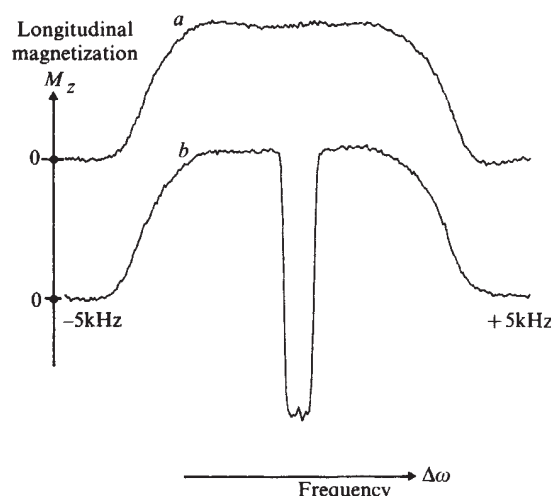


Fig. 2 The magnetization profile (a) created by applying a field gradient down the axis of a tube of water whose length was greater than that of the receiving coils. When the selective π pulse of Fig. 1 was applied, a sharp hole was punched in the middle of the profile (b).

changed. For values of $\mu \geq 4$, the width of the selected region is given by

$$\delta\omega \approx \pm\mu\beta \quad (8)$$

when the driving function has an amplitude

$$\Omega_0 \gg \mu\beta \quad (9)$$

However, the value $\mu \approx 5$ seems to give the most selective result.

Experimental verification was achieved by setting $\beta = 400 \text{ s}^{-1}$, and creating, with the aid of an array processor and twin D/A converters, the necessary h.s. waveforms of duration 25 ms; see Fig. 1. These were then applied to the single sideband modulator which is part of our spectrometer's transmitter circuitry, to create the necessary irradiating field. The sample used was a 30-cm long tube of water placed in a solenoidal probe coil of length 8 cm and diameter 2.5 cm. An axial field gradient equivalent to 1 kHz cm^{-1} was then applied; the magnetization profile thus created is shown in Fig. 2a. The selective h.s. pulse was given and the magnetization interrogated with the aid of an ordinary 90° pulse, followed by data accumulation and Fourier transformation in the usual manner. Figure 2b shows the rectangular hole punched in the frequency profile. The sides are exceedingly sharp. While the width of the inversion is given by $\mu\beta/\pi \approx 640 \text{ Hz}$ (compare with $\Omega_0 \gg \mu\beta$), the sides of the hole cover only

~60 Hz each, as might be expected from a 25-ms pulse. Increasing the incident radiofrequency power by a factor of up to 500 above the minimum required for inversion had negligible effect on the profile. Further increase distorted the profile as the r.f. power amplifier became nonlinear. While the magnetization in the selected region is highly perturbed and finally inverted, magnetization outside that region must also execute a considerable motion in the presence of high-power pulses. Although we do not yet know the trajectories followed during the pulse by various magnetizations, the experimental results unequivocally indicate that magnetization always comes to rest upright outside the region and inverted within it—a unique phenomenon as far as we know.

Note added in proof: A simple closed form of equation (5) has been found and a description is being prepared for publication.

Received 9 April; accepted 20 June 1984.

1. McCall, S. L. & Hahn, E. L. *Phys. Rev.* **183**, 457–485 (1969).

2. Lamb, G. L. *Phys. Rev. A* **9**, 422–430 (1974).

A focal brain abnormality in panic disorder, a severe form of anxiety

Eric M. Reiman*, Marcus E. Raichle†‡, F. Kevin Butler*, Peter Herscovitch†‡ & Eli Robins*

* Departments of Psychiatry, † Neurology and Neurological Surgery, and ‡ Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, Missouri 63110, USA

Panic disorder is characterized by recurrent anxiety attacks in the absence of a frightening stimulus¹. It is a common disorder, affecting 2–5% of the general population and 10–14% of patients seen in cardiology practice^{2,3}. Infusion of sodium (DL)lactate precipitates an anxiety attack in most persons with this disorder but rarely does so in normal controls, suggesting a neurobiological basis for the problem^{4–6}. Despite this observation, the pathophysiology of panic disorder remains unknown. We have now used positron emission tomography to measure cerebral blood flow (CBF) in patients with panic disorder in the absence of a panic attack. Analysis of CBF in regions thought to mediate symptoms of panic, anxiety and vigilance reveals a significant ($P < 0.005$) abnormal asymmetry of CBF (left < right) located in a region of the parahippocampal gyrus. This asymmetry was present in seven patients with panic disorder and a positive response to lactate infusion but was absent in six normal controls and in three patients with panic disorder associated with a negative response to lactate. We believe this to be the first study to identify a discrete brain abnormality in patients with this severe form of anxiety.

Ten patients with a history of panic attacks and six normal control subjects were studied. All patients satisfied both DSM III criteria¹ and Feighner criteria⁷ for panic disorder. They had no other psychiatric disorders and were physically well. Measurements of regional CBF were made in each subject using the PETT VI system^{8,9} which simultaneously records data from seven slices with a centre-to-centre separation of 14.4 mm. All studies were done with an in-plane resolution of 11.7 mm full width at half maximum in the centre of the field of view. Regional CBF was measured by the intravenous injection of ⁵O-labelled water using a modification of the Kety autoradiographic model developed and validated in our laboratory^{10,11}. All subjects were studied in the resting state. During the scans the subjects lay quietly with eyes closed and received minimal sensory stimulation. Ambient noise consisted of cooling fans for electronic equipment. All of the patients and two of the controls subsequently underwent a lactate challenge⁴ during which 10 mg per kg of 500 mM sodium (DL)lactate was infused over 20–

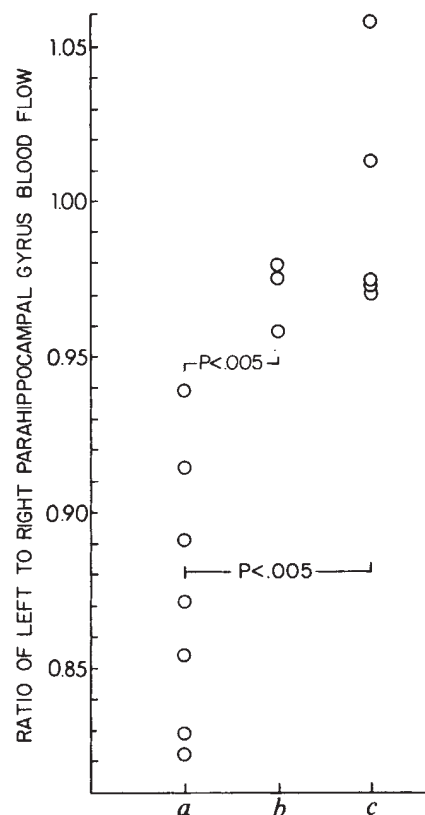


Fig. 1 Left-to-right ratio of blood flow in the parahippocampal gyrus of: *a*, patients with a history of panic disorder and a positive response to sodium lactate infusion (that is, a panic attack⁵); *b*, patients with a history of panic disorder and negative response to sodium lactate infusion; *c*, normal control subjects. These data were obtained in a quiet, resting state.

Table 1 Location of selected cerebral regions in a stereotactic atlas¹⁹

	X (cm)	Y (cm)	Z (cm)
Parahippocampal gyrus	±2.3	-2.3	-0.3
Hippocampus	±3.0	-1.0	-1.5
Hypothalamus	±0.7	0.8	-0.5
Orbito-insular gyri	±3.4	2.7	-0.7
Anterior cingulate gyrus	±0.7	4.0	2.0
Amygdala	±2.3	0.9	-1.7
Inferior parietal lobule	±3.0	-3.8	4.0

X, Lateral distance to the left (positive) or to the right (negative) of the centre of the brain slice containing the region of interest (ROI). Y, Distance anterior (positive) or posterior (negative) to the centre of the brain slice containing the ROI. Z, Vertical distance of the brain slice containing the ROI above (positive) or below (negative) a horizontal plane through the anterior and posterior commissures.

30 min. These subjects were aware that lactate would be infused at an undisclosed time and that this procedure might precipitate a panic attack.

We divided our subjects into three groups: group *a*, seven patients with a history of panic disorder and a positive response to lactate infusion (that is, a panic attack as defined by Pitts and McClure⁴); group *b*, three patients with panic disorder and a negative response to lactate infusion; group *c*, six normal controls who had no history of psychiatric disorders and who were physically well. Whole brain, hemisphere and regional CBF were computed from our data in all subjects. The regional measurements were obtained bilaterally in seven areas of brain thought to mediate symptoms of panic, anxiety and vigilance^{12–16}: hippocampus, parahippocampal gyrus, inferior parietal lobule, anterior cingulate gyrus, hypothalamus, orbito-

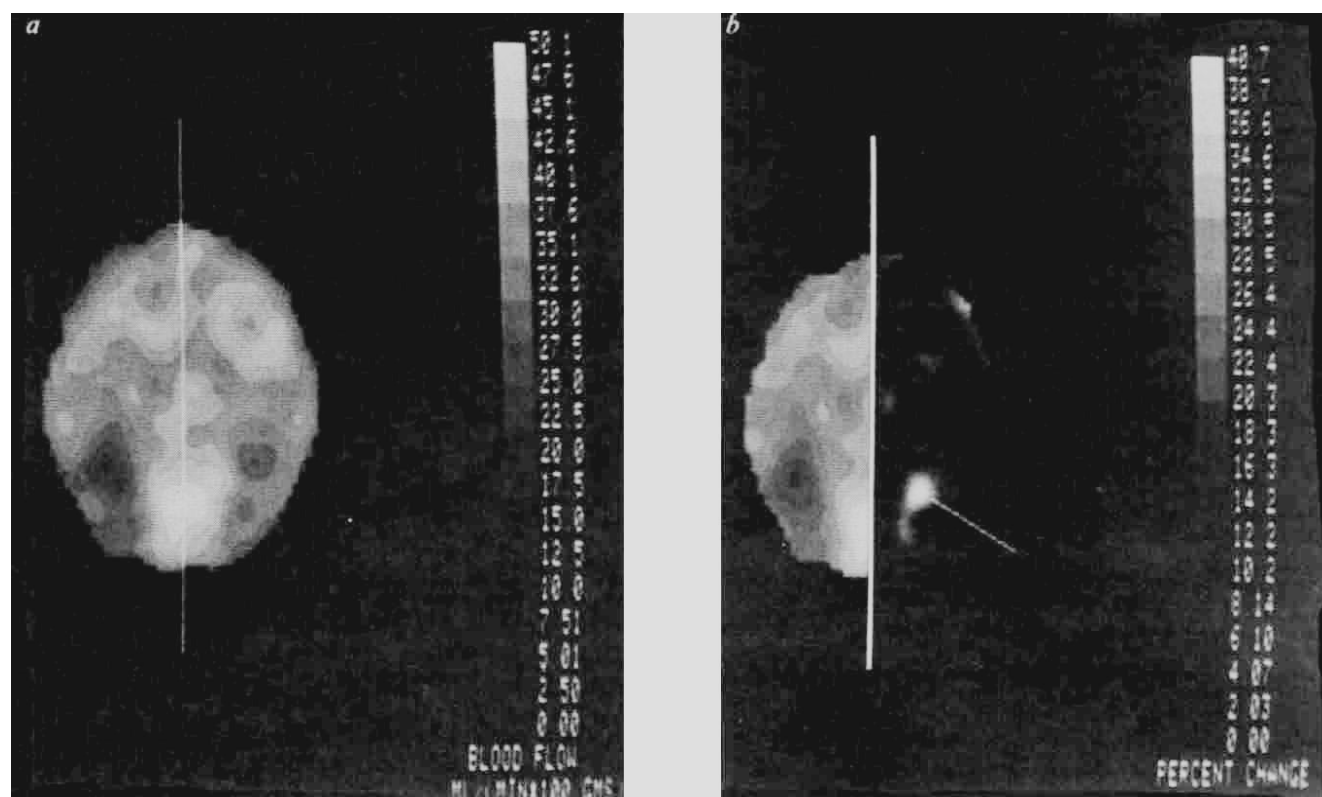


Fig. 2 Positron emission tomographic images of regional cerebral blood flow (CBF) obtained in resting conditions in a patient with panic disorder and a positive response to lactate. For each image the patient's left is at 9 o'clock and anterior is at 12 o'clock. Image *a* is the horizontal section containing the parahippocampal region. The image is a quantitative representation of cerebral blood flow corresponding to the calibrated grey scale. Image *b* is the same tomographic section with the left half representing cerebral blood flow and the right half representing the difference in blood flow between right and left hemispheres, expressed as a percentage of the left hemisphere blood flow in the grey scale. All asymmetries depicted on the image are right > left. The region of abnormal asymmetry corresponds to the area of maximal asymmetry indicated by the line at the right of the image. This parahippocampal asymmetry was observed in all patients in group *a* (panic disorder with a positive response to sodium lactate), but not in the other subjects (see Fig. 1). Other asymmetries appearing in the image varied randomly among patients.

insular gyri and amygdala. These regions of interest were identified in a stereotactic atlas¹⁷ of the brain. Their locations on the positron emission tomography images were then found using an anatomical localization technique that is entirely independent of the visual appearance of these images¹⁸. All regions measured 13.5×13.5 mm in the plane of the slice and were 13.9 mm thick. The regions and their atlas coordinates are listed in Table 1.

The data were analysed using one-way analysis of variance techniques, including the Newman-Keuls multiple range test where appropriate. The three subject groups did not differ significantly in CBF in whole brain, left hemisphere or right hemisphere or in the left-right ratio of hemisphere CBF. We then compared the left to right ratios for CBF in seven preselected regions (Table 1) as a highly sensitive index of regional asymmetries. This analysis revealed a significantly lower ratio in the parahippocampal gyrus of the seven patients in group *a* who had a positive lactate infusion response (Table 2, Fig. 1) than in groups *b* or *c* ($P < 0.005$). Figure 2 shows a typical study.

We were unable to determine whether this asymmetry represented a decrease in the left parahippocampal gyrus or an increase in the right parahippocampal gyrus because the absolute values for CBF fall within our range of normal values. No other abnormal asymmetries were found. There was no group effect for age, sex, handedness or medication.

After completion of this study, we analysed data from 20 additional neurologically normal volunteers. Psychiatric histories were not available at the time of the analysis. Ratios of parahippocampal blood flow were indistinguishable from those of our six original controls, with one exception, a 34-yr-old female who had a ratio of 0.87; this is the mean for our seven patients with panic disorder and a positive response to lactate (Table 2, Fig. 1). This subject was interviewed and underwent a lactate challenge. Her history was consistent with the diagnosis of panic disorder¹ and she had a positive response to intravenous sodium lactate. The other 19 subjects were subsequently interviewed and none gave a history of psychiatric illness.

Table 2 Left-to-right ratio (mean $\pm$ s.d.) of cerebral blood flow in selected regions in various human subjects

Group	Parahippocampus	Hippocampus	Inferior parietal lobule	Anterior cingulate gyrus	Hypothalamus	Orbito-insular gyri	Amygdala
<i>a</i> ($n=7$)	0.87 ± 0.04	1.03 ± 0.06	1.01 ± 0.09	0.88 ± 0.15	1.04 ± 0.12	1.06 ± 0.10	1.01 ± 0.12
<i>b</i> ($n=3$)	0.97 ± 0.01	0.97 ± 0.10	1.01 ± 0.04	0.80 ± 0.07	0.99 ± 0.08	0.91 ± 0.08	1.12 ± 0.33
<i>c</i> ($n=6$)	1.00 ± 0.03	1.00 ± 0.04	1.03 ± 0.11	0.82 ± 0.11	0.99 ± 0.05	0.95 ± 0.08	1.04 ± 0.10
$F(d.f.=2, 13)$	20.41*	0.86	0.10	0.54	0.50	3.28	0.53

Group *a*, patients with panic disorder and a positive response to lactate infusion; group *b*, patients with panic disorder and a negative response to lactate infusion; group *c*, normal controls. The exact locations of the brain regions studied are given in Table 1.

* Significant at $P < 0.0001$.

Thus, we have demonstrated here a focal brain abnormality in patients with a discrete psychiatric disorder. Seven patients with a history of both spontaneous and lactate-induced panic attacks were distinguished from other patients with panic attacks and six normal controls by asymmetry in local CBF (and, by inference, local metabolism¹⁹) in a circumscribed region of the parahippocampal gyrus. This abnormality seems to correlate with vulnerability to panic attacks but not with the panic state itself.

The abnormality identified here may reflect an exaggeration of a normal hemisphere specialization for the expression of emotional responses. It is located in a region of the parahippocampal gyrus that contains both entorhinal cortex and subiculum, the major input and output, respectively, of the hippocampus. Thus, a predisposition to panic attacks is associated with a discrete abnormality in a region of the human brain thought to be important in the expression of emotion¹³.

This work was supported by USPHS grants NS06833, HL13851, RR00036 and the McDonnell Center for the Studies of Higher Brain Function. We thank Drs Joel Perlmutter, Peter Fox, Cliff Saper, William Powers and Warren Danziger for their helpful suggestions and critical comments, and the technical staff of the Division of Radiation Sciences.

Received 12 March; accepted 12 June 1984.

1. *Diagnostic and Statistical Manual Vol. 3* (American Psychiatric Association, Washington, D.C., 1980).
2. Sheehan, D. V. *New Engl. J. Med.* **307**, 156–158 (1982).
3. Brown, J. T., Murlow, C. D. & Stoudemire, G. A. *Ann. intern. Med.* **100**, 558–564 (1984).
4. Pitts, F. N. & McClure, J. N. Jr *New Engl. J. Med.* **277**, 1329–1336 (1967).
5. Fink, M., Taylor, M. A. & Vocavka, J. *New Engl. J. Med.* **281**, 1429–1440 (1970).
6. Ackerman, S. H. & Sachar, E. J. *Psychosom. Med.* **36**, 69–81 (1974).
7. Feighner, J. P. et al. *Archs gen. Psychiat.* **26**, 57–63 (1972).
8. Ter-Pogossian, M. M., Ficke, D. C., Hood, J. T., Yamamoto, M. & Mullani, N. A. *J. Comput. Assisted Tomogr.* **6**, 125–133 (1982).
9. Yamamoto, M., Ficke, D. C. & Ter-Pogossian, M. M. *IEEE Trans. nucl. Sci.* **NS29**, 529–533 (1982).
10. Herscovitch, P., Markham, J. & Raichle, M. E. *J. nucl. Med.* **24**, 782–789 (1983).
11. Raichle, M. E., Martin, W. R. W., Herscovitch, P., Mintum, M. A. & Markham, J. *J. nucl. Med.* **24**, 790–798 (1983).
12. Papez, J. W. *Archs Neurol. Psychiat.* **38**, 725–743 (1937).
13. Gray, J. A. G. *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System* (Oxford, New York, 1982).
14. Saper, C. B. *J. comp. Neurol.* **210**, 163–173 (1982).
15. Heilman, K. M. & Valenstein, E. *Clinical Neuropsychology* (Oxford, New York, 1979).
16. Mesulam, M. M. *Ann. Neurol.* **10**, 309–325 (1981).
17. Talairach, J. & Szikla, G. *Atlas D-Anatomie Stereotaxique du Telencephale* (Masson & Cie, Paris, 1967).
18. Fox, P. T., Perlmutter, J. S. & Raichle, M. E. *J. Comput. Assisted Tomogr.* **9**(1) (in the press).
19. Raichle, M. E., Grubb, R. L. Jr, Gado, M. H., Eichling, J. O. & Ter-Pogossian, M. M. *Archs Neurol.* **33**, 523–526 (1976).

Initiation of synchronized neuronal bursting in neocortex

Barry W. Connors

Department of Neurology, Stanford University School of Medicine, Stanford, California 94305, USA

Epilepsy is characterized by highly synchronized paroxysmal bursts of activity within a large population of cortical neurones^{1,2}. Because such spontaneous, synchronized discharges can occur even in isolated blocks of neocortex³, mechanisms for initiating and coordinating this activity must reside within the cortex itself. However, the specific cellular properties and local neural circuitry responsible for such behaviour are unknown. In a previous study of neocortex *in vitro*, we found that treatment with the convulsants penicillin and bicuculline led to synchronized bursts which were driven by unusually large and long-lasting excitatory synaptic conductances⁴. I now report evidence that synchronized bursts are initiated by a small, spacially discrete subpopulation of cells located in the area comprising layer IV and upper layer V. Neural elements in these layers appear to project paroxysmal synaptic excitation radially, onto the neurones of other layers.

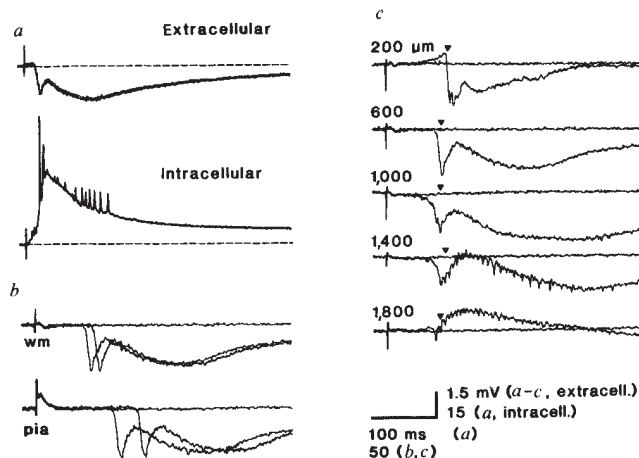
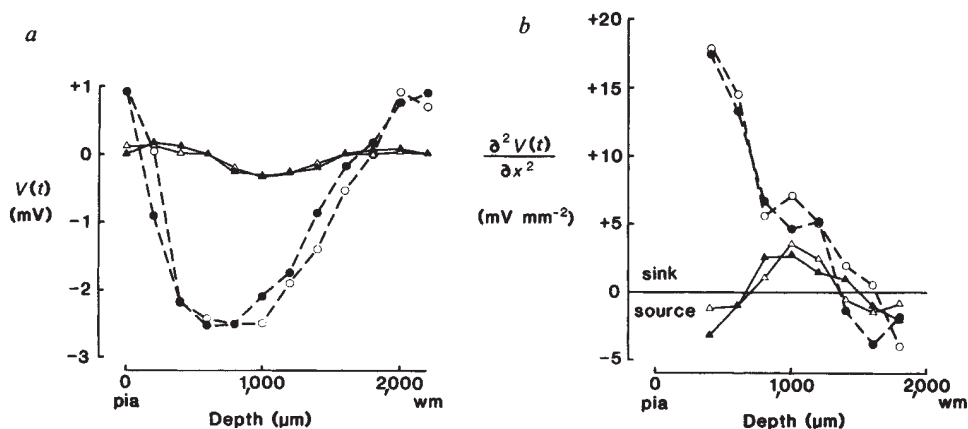


Fig. 1 Synchronized bursting in neocortical slices treated with 50 μ M bicuculline. *a*, Stimulation of the white matter (vertical artefact) elicited prolonged neuronal burst, recorded intracellularly (lower), and a simultaneous paroxysmal field potential (p.f.p.; upper). Voltage records in all figures are negative downward. Intracellular recording was obtained 1,000 μ m from the pia, extracellular recording 600 μ m from the pia. *b*, Different pathways of activation evoke similar synchronized bursts. P.f.p.s were recorded 600 μ m from the pia and single stimuli were applied to either white matter (wm) or pia. Intensity was adjusted to be near threshold and shocks were applied every 4 s. Each set of traces shows three consecutive stimuli, two of which evoked p.f.p.s and one of which failed to evoke a p.f.p. Threshold intensity for pial activation was about five times that from white matter. *c*, p.f.p.s recorded from different laminae. Conditions were the same as in *b*; threshold stimuli were delivered to the white matter and two consecutive traces are superimposed at each location. Distances were measured from the pia. Each location was examined in sequence by a moveable electrode; a second electrode was located at 800 μ m from the pia to monitor shifts in p.f.p. latency (not shown). Arrowheads on each set of traces identify the latency of the first negative peak recorded at 800 μ m.

Methods: Slices were continuously superfused with physiological solution at 35 °C. Intracellular recordings were made with 4 M K acetate-filled micropipettes (30–40 M Ω); extracellular field potentials were measured with 1 M NaCl-filled pipettes (3–7 M Ω). Monopolar cathodal stimuli (200 μ s duration) were delivered through sharpened tungsten wires. Data were recorded on magnetic tape, digitized and analysed off-line.

Experiments were performed on isolated, coronally sectioned slices (400–500 μ m thick) of sensorimotor neocortex from mature guinea pigs, as described previously⁵. Synchronized bursting was induced by bathing the slices in solutions containing 1–50 μ M bicuculline methiodide, a potent antagonist of the cortical inhibitory neurotransmitter γ -aminobutyric acid (GABA)⁶. Spontaneous bursting was present only at the higher bicuculline concentrations. Low-intensity electrical stimuli applied to the subcortical white matter, however, could elicit essentially all-or-none neuronal bursts of long and variable latency at all dosages (Fig. 1*a*). Blockage of chemical synaptic transmission, by reducing calcium concentration from 2 to 1 mM and adding 3 mM manganese, abolished all bursting. The bursting induced by disinhibition seems to be synchronized among a very high proportion of the cortical neurones. When the activity of pairs of neurones was recorded simultaneously, bursts were invariably phase-locked⁴. This was reflected in the extracellular voltage; each synchronized burst was concurrent with a paroxysmal field potential (p.f.p.) that represented the simultaneous activity of a large number of neurones (Fig. 1*a*). P.f.p.s did not occur in the absence of intracellularly recorded bursts. The p.f.p.s displayed several traits identical to the intracellularly recorded bursts, including long and very variable latencies, all-or-none occurrence and an invariant shape when activated by different afferent pathways (Fig. 1*b*).

Fig. 2 Laminar analysis of extracellular field potentials during synchronized bursts. Data are taken from same experiment as illustrated in Fig. 1c. *a*, p.f.p. amplitudes at two different latencies ($V(t)$) as a function of sub-pial depth. Circles plot measurements at the latencies of the first negative peak at 800 μm (that is, arrowheads in Fig. 1c). Triangles represent data measured 12.5 ms before this point, that is, the earliest part of the p.f.p. which could be easily resolved from the background noise. Filled symbols were derived from p.f.ps initiated by stimulation of white matter, open symbols from pia stimuli. Five sweeps were averaged to derive measurements at each point. *b*, Current-source densities calculated from data shown in *a*. Net membrane currents were estimated by numerically approximating the second spatial derivative of $V(t)$ ($\partial^2 V(t)/\partial x^2$) using a finite difference formula⁸ with a spacing of 400 μm . Current sinks are indicated by values greater than zero; negative values imply current sources. Cortical resistivity along the axis of measurement was assumed constant¹⁸.



To determine the initial site of active membrane current during synchronized bursts, p.f.ps were recorded at defined distances between the pia and white matter (Figs 1c, 2a). An activating stimulus was kept near threshold so that non-burst-related potentials could be subtracted, and variations in latency were monitored by a second electrode fixed 800 μm from the pia. A one-dimensional current-source density analysis was performed by estimating the second spatial derivative of the extracellular voltage at various latencies⁷⁻⁹. The earliest burst-related membrane current sink was invariably centred 900–1,100 μm from the pia ($n=5$ slices; Fig. 2b). As the burst progressed, the maximal sink current increased greatly in size and moved superficially, to within 400–600 μm of the pia. Within a slice, p.f.ps and their associated current-source density distributions were not dependent on the site (that is, pia or white matter) of the electrical stimulus (Fig. 2). These data suggest that synchronized bursts are invariably presaged by active inward current in a subpopulation of middle-layer neuronal elements.

Electrical stimulation of the slice elicits a relatively undefined pattern of activation. Therefore, I attempted to initiate synchronized bursts by selectively exciting small groups of neocortical neurones using focally applied L-glutamate, an excitatory amino acid, or focally elevated potassium concentration. When a threshold level of glutamate was exceeded, typical synchronized bursts were elicited (Fig. 3a); this threshold varied systematically across laminae (Fig. 3b). Sharp threshold minima always occurred in the middle layers, peaking at a distance of $965 \pm 225 \mu\text{m}$ (mean $\pm$ s.d.) from the pia ($n=12$ slices). When tested in the same slice, K^+ ejections (25 mM) yielded the same site of threshold minimum ($n=3$ slices). When the loci of glutamate threshold minima were marked with small spots of pontamine blue dye and examined histologically, they were found to lie within layer IV or the upper part of layer V (data not shown).

Several studies of neocortex *in situ*, using various species and cortical areas, have shown that specific middle layers are especially sensitive to the convulsant effects of penicillin¹⁰⁻¹². Similar results were obtained here with bicuculline and neocortex *in vitro* ($n=6$ slices). Slices were bathed in convulsant-free solutions and focal pressure applications of the drug were made at varying distances from the pia (Fig. 4a). When a suitably small bicuculline pulse was applied to superficial ($<800 \mu\text{m}$ from the pia) or deep ($>1,200 \mu\text{m}$ from the pia) cortical layers, evoked potentials were either unaffected or, at the recording site closest to the application, slightly increased in amplitude (Fig. 4b, 640 and 1,500 μm). However, the same amounts of bicuculline ejected into the middle layers often led to the generation of relatively small, all-or-none field potentials of long and shifting latency (Fig. 4b, 1,100 μm). Coincident with each of these events, intracellular recordings ($n=12$) from neurones

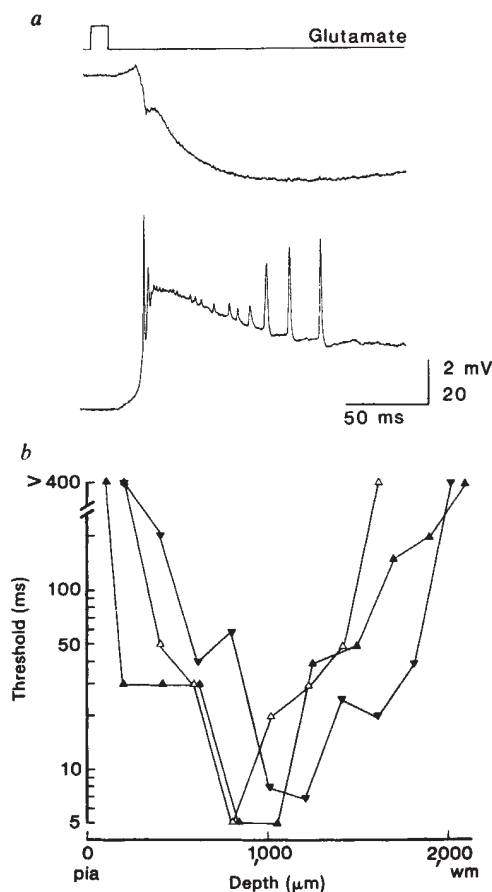
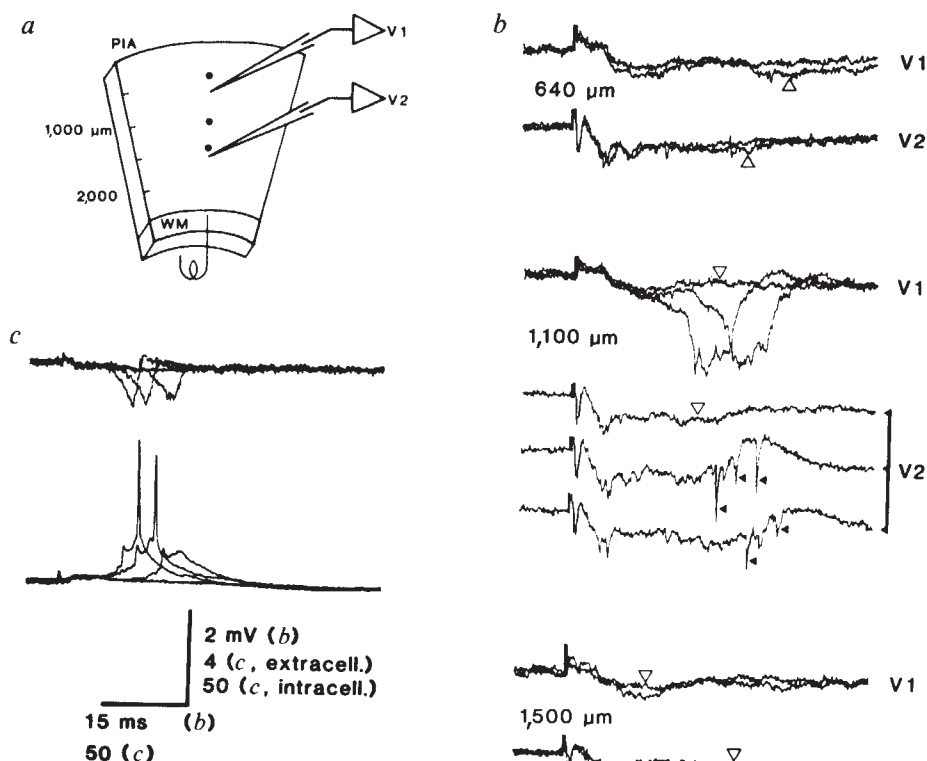


Fig. 3 Triggering of synchronized bursts by L-glutamate. *a*, L-Glutamate applied 1,000 μm from the pia (top) evoked a p.f.p. (middle) and a simultaneous neuronal burst, recorded intracellularly (bottom). Extracellular recording was made 500 μm from the pia, intracellular recording 700 μm from the pia. *b*, Threshold for p.f.p. generation (in ms of L-glutamate pulse duration) as a function of the distance of glutamate pipette from the pia. Examples from three different slices are illustrated. Each point signifies the pulse duration at which p.f.ps were generated on 50% of the trials.

Methods: L-Glutamate (0.5–2 mM in physiological solution) was pressure ejected (3.5 kg per cm^2 of N_2) through micropipettes of 2–5 μm tip diameter. The duration of the pressure pulses was varied by digital control of a rapidly switching solenoid valve (General Valve Inc.). A minimum command of 5 ms was required to open the valve.

Fig. 4 Laminar sensitivity to focally applied bicuculline. The tissue slice was bathed in bicuculline-free solution. *a*, Schematic diagram of experimental design. Low-intensity shocks were applied to the white matter (0.1 Hz) and two extracellular electrodes (V1 and V2) were located at fixed distances of 700 μm and 1,600 μm from the pia. Bicuculline was released in constant amounts from a micropressure pipette (500 ms pulse; 100 μM in physiological solution) at three distances from the pia (dots). *b*, Only middle-layer ejections led to paroxysmal potentials. Each set of traces shows records from V1 and V2 before (arrowheads) and 1 min after bicuculline ejections. Ejections were spaced 10 min apart; each location was tested twice and similar results were observed. Only the ejection at 1,100 μm led to large, negative, shifting latency waves at V1. The simultaneous response at V2 was smaller and positive-going, but extracellular action potentials (filled triangles) were seen to discharge synchronously with the shifting late wave at V1. V2 traces at 1,100 μm are separated for clarity. *c*, Intracellular recordings (bottom) confirm that neurones displayed paroxysmal depolarizations during extracellular negativities (top). Four consecutive superimposed traces are shown, including three paroxysmal potentials and one failure. Recordings were made from same slice as for *b*.



near the extracellular recording sites revealed a depolarizing wave (8–25 mV, 25–65 ms) which could evoke from one to four action potentials (Fig. 4c). Such behaviour is very similar to the paroxysmal synchronized bursting observed in slices uniformly exposed to penicillin⁴, a weaker convulsant agent.

These experiments provide three converging lines of evidence which implicate a specific region of neocortex in the initiation of synchronized bursts. The area comprising layer IV and superficial layer V is where: (1) the first burst-related inward membrane current occurs, (2) the minimal amount of L-glutamate excitation evokes a burst, and (3) relatively small areas of disinhibition lead to widespread synchronized bursts. Further, *in vivo* studies of cortical epileptic foci in rats have shown that reductions in extracellular calcium activity precede burst onset most prominently in the middle layers¹³. These data are most simply explained by assuming the presence of a subset of middle-layer neurones with properties unique within the neocortex⁴. Such cells would have membranes endowing them with intrinsic bursting capabilities, a system of short-latency reciprocal excitation to synchronize their activity, an extensive set of divergent axons to powerfully excite other cortical neurones within a local area, and adjacent local circuits of GABA-mediated inhibition to control bursting in normal conditions. The middle layers pinpointed by this investigation coincide with the position of a small population of intrinsically bursting neocortical neurones observed in control recording conditions⁵. These cells must now be considered strong candidates for a pacemaking function in synchronized discharge. Within hippocampus, pyramidal cells of the CA2/CA3 region apparently serve an analogous role by virtue of their specific membrane properties and excitatory synaptic contacts^{14,15}.

The possible task of a middle-layer synchronizing system in normal neocortical function is unknown. An oscillatory synaptic drive from these neurones may underlie the spontaneous

spindle-like rhythms of extracellular voltage often observed in large, isolated neocortical slabs *in situ*^{3,16}. Additionally, given their position, such cells probably receive direct monosynaptic contacts from specific thalamic afferents¹⁷, implying an important role in the integration of sensory information. Thus, the temporal patterns of both normal and epileptic cerebral activity may result, in part, from the interactions of separate cortical and subcortical (that is, thalamic) neural pacemakers¹⁶.

I thank C. P. George for help in some of the early experiments, Drs D. A. Prince, B. R. Ransom and J. W. Lighthall for their comments on the manuscript, S. Lyon and J. L. Kadis for technical assistance and C. A. Joo for secretarial help. This study was supported by Biomedical Research Support grant RR 5353 and NS 12151 from the NIH.

Received 4 January; accepted 4 June 1984.

1. Matsumoto, H. & Ajmone-Marsan, C. *Expl Neurol.* **9**, 286–304 (1964).
2. Jasper, H. H., Ward, A. A. & Pope, A. in *Basic Mechanisms of the Epilepsies* (Little, Brown, Boston, 1969).
3. Kristiansen, R. & Courtois, G. *EEG clin. Neurophysiol.* **1**, 265–272 (1949).
4. Gutnick, M. J., Connors, B. W. & Prince, D. A. *J. Neurophysiol.* **48**, 1321–1335 (1982).
5. Connors, B. W., Gutnick, M. J. & Prince, D. A. *J. Neurophysiol.* **48**, 1302–1320 (1982).
6. Johnston, G. A. R. *Rev. Pharmac. Tox.* **18**, 269–289 (1978).
7. Nicholson, C. & Freeman, J. A. *J. Neurophysiol.* **38**, 356–365 (1975).
8. Freeman, J. A. & Nicholson, C. *J. Neurophysiol.* **38**, 369–382 (1975).
9. Mitzdorf, U. & Singer, W. *Expl Brain Res.* **33**, 371–394 (1978).
10. Lockton, J. W. & Holmes, O. *Brain Res.* **190**, 301–304 (1980).
11. Ebersole, J. S. & Chatt, A. B. *Neurology* **30**, 355 (1980).
12. Elger, C. E., Speckmann, E. J., Prohaska, O. & Caspers, H. *EEG clin. Neurophysiol.* **51**, 393–402 (1981).
13. Pumain, R., Kurcewicz, I. & Louvel, J. *Science* **222**, 177–179 (1983).
14. Schwartzkroin, P. A. & Prince, D. A. *Brain Res.* **147**, 117–130 (1978).
15. Wong, R. K. S. & Traub, R. D. *J. Neurophysiol.* **49**, 442–458 (1983).
16. Andersen, P. & Andersson, S. A. in *Physiological Basis of the Alpha Rhythm* 71–83 (Appleton-Century-Crofts, New York, 1968).
17. White, E. L. *Brain Res. Rev.* **1**, 275–312 (1979).
18. Mitzdorf, U., Appendix to Mitzdorf, U. & Singer, W. *J. Physiol., Lond.* **304**, 203–220 (1980).

Inducible expression of H-2 and Ia antigens on brain cells

Grace H. W. Wong*, Perry F. Bartlett, Ian Clark-Lewis, Frank Battye & John W. Schrader

The Walter and Eliza Hall Institute of Medical Research, Royal Melbourne Hospital, Victoria 3050, Australia

Cells in the brain express unusually low levels of antigens encoded by the major histocompatibility complex (MHC)^{1,2}. This is somewhat surprising as class I (H-2) and class II (Ia) MHC antigens have critical roles in immune responses³. The activation of T lymphocytes is associated with the enhanced expression of these antigens and this effect is mediated by a specific T-cell lymphokine, γ -interferon (IFN- γ)⁴⁻¹⁷. Here we show that IFN- γ induces a dramatic increase in the expression of H-2 antigens on the cells of the brain. After exposure to IFN- γ *in vitro*, all surviving cells, including most astrocytes, oligodendrocytes, microglia and at least some neurones, express H-2 antigens. Direct injection of IFN- γ into the brains of mice indicated that H-2 antigens were also induced *in vivo*. Furthermore, IFN- γ induced Ia antigens on a subpopulation of astrocytes. The induction of H-2 antigens by IFN- γ may render brain cells competent to initiate and participate in immune reactions and may therefore contribute to both immunoprotective and immunopathological responses in the brain.

Cell suspensions were prepared by trypsinization of brains from 2-day-old mice. H-2 antigens were detected by monoclonal antibodies and indirect immunofluorescence and the levels quantitated by flow cytometry. The proportion of cells expressing detectable H-2 antigens in the starting population was less than 1%. These cells were cultured with or without IFN- γ produced by monkey (COS-I) cells that had been transfected with the cloned murine IFN- γ gene¹⁸. Within 24 h of IFN- γ treatment, H-2 antigens were detectable on 99% of viable cells (Fig. 1) and 87% of cells expressed levels of H-2 antigens similar to those on IFN- γ -treated macrophages or spleen cells. The mean H-2-specific fluorescence per cell was increased by at least 60-fold. Brain cells exposed either to control supernatants containing no IFN- γ (Fig. 1) or to medium alone did not express significantly higher levels of H-2 antigens than did the starting population.

Detectable induction of H-2 antigens required an 18-h exposure to IFN- γ and maximal induction occurred at 24–48 h. Induction was dose-dependent and the maximal response was observed at a concentration of IFN- γ corresponding to 1–2 units per ml of antiviral activity. Similar results were obtained with IFN- γ derived from antigen-stimulated T-cell clones or mitogen-activated spleen cells. Actinomycin D (0.05 μ g ml⁻¹) and cycloheximide (5 μ M) blocked the induction of H-2 antigens without affecting cell viability, suggesting that the induction required new RNA and protein synthesis. The induction of H-2 antigens is not due to inhibition of cell division because mitomycin C (40 μ g ml⁻¹) inhibited DNA synthesis but did not induce H-2 antigens on brain cells.

The brain cell types that expressed H-2 antigens were identified by antibodies to cell type-specific components, glial fibrillary acidic protein (GFAP) for astrocytes¹⁹, galactocerebroside (GC) for oligodendrocytes^{20,21}, Fc receptors for microglia²² and by a monoclonal antibody A2B5 for neurones²³. Although it has been shown that A2B5 also binds to glial progenitor cells²⁴, the A2B5⁺ cells we examined were GFAP⁺ and GC⁺, suggesting that they were neurones. Neurones were also identified with an antiserum that reacts with neurofilaments (NF)²⁵. The cellular composition of the viable cell population before and after 24 h in culture with or without IFN- γ is shown in Table 1. Double-

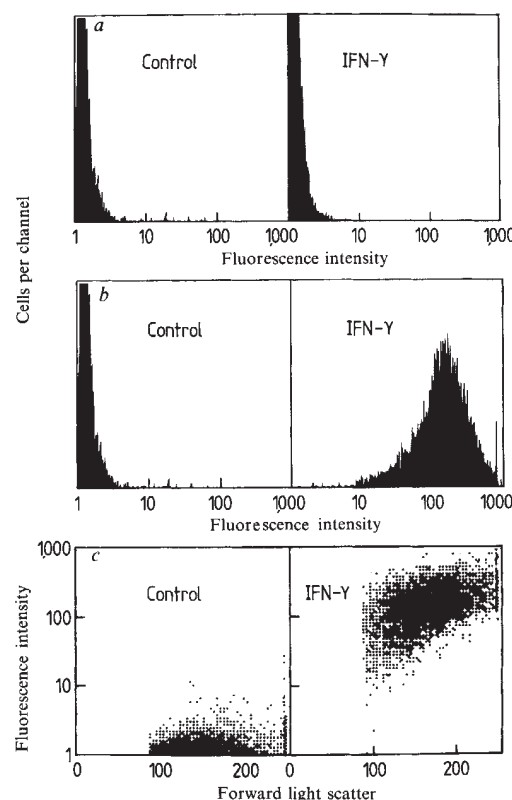


Fig. 1 Induction of H-2 antigens on brain cells. *a*, Analysis of cells stained with fluorescein-conjugated avidin alone; *b*, cells stained with biotin-conjugated monoclonal antibody (11-4.1) to H-2K^k antigens and fluorescein-conjugated avidin. *c*, Results of analysis of the same samples as in *b* except that fluorescence intensity (H-2 antigens) is plotted against forward light scatter, an indicator of cell size. The level of increase in H-2 antigens was calculated as the ratio of mean fluorescence intensity of the IFN- γ -treated sample to that of the control.

Methods: Cell suspensions were prepared from entire brains of 2-day-old CBA mice. The tissue was carefully dissected free from the meninges, finely sliced and incubated at 37 °C for 30 min with trypsin (1 mg ml⁻¹) and DNase I (10 μ g ml⁻¹) to prevent the formation of aggregates. Trypsinization was arrested by the addition of fetal calf serum (10%). The cell suspension was centrifuged and the pellet was dispersed by pipetting. Approximately 1.2×10^7 viable cells were recovered per brain and the viability was ~80%. These cells (~ 10^6 per ml) were cultured in Dulbecco's modified Eagle's medium with fetal calf serum (10%) in Corning flasks (75 cm²) at 37 °C with either supernatant (5%) of cultures of COS-I cells transfected with the cloned murine IFN- γ gene¹⁸ containing 40 units per ml of IFN- γ or a control supernatant (5%) of cultures of nontransfected COS-I cells (both supernatants were a gift from P. Gray and D. Goeddel) (Genentech). IFN- γ was assayed as previously described⁴. After 24 h, culture of the recovery of total viable cells was 38% and 35% in the absence and presence of IFN- γ (2 units per ml), respectively. The non-adherent cells and adherent cells removed by trypsinization were pooled and stained with optimal amounts of biotin-conjugated mouse monoclonal antibody against H-2K^k antigens (11-4.1)^{5,6} at 4 °C for 20 min, followed by washing and addition of fluorescein-conjugated avidin (Sigma). Trypsinization has been reported to have no effect on the expression of MHC antigens³³⁻³⁵. Detection of the observed fluorescence was specific for H-2 antigens in that a monoclonal antibody (34-1-2S)^{5,6} of the same class (IgG2a), but specific for a different H-2 haplotype ((H-2K^dD^d), did not bind to IFN- γ -treated brain cells from CBA (H-2K^k, D^k) mice. Approximately 10,000 stained cells per sample were analysed by flow cytometry in a fluorescence-activated cell sorter (FACS II, Becton-Dickinson). The % of the viable cells within the set gates were 90 and 91% of the control and IFN- γ -treated samples, respectively. The intensity of fluorescence (proportional to the level of H-2 antigens) is shown in arbitrary units on a three-decade logarithmic scale. The non-viable cells in the sample were stained with propidium iodide to allow their exclusion from analysis as cells that fluoresce bright red.

* Present address: Department of Microbiology, University of Southern California, School of Medicine, Los Angeles, California 90033, USA.

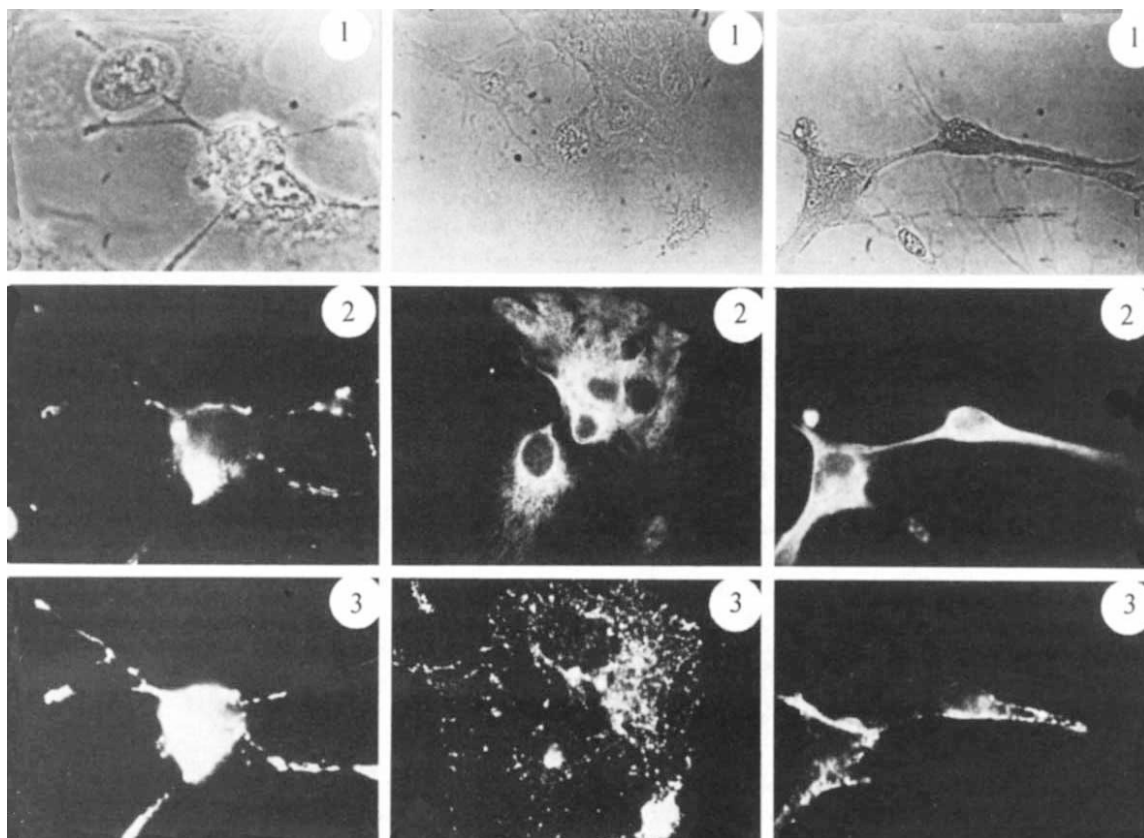


Fig. 2 Induction of H-2 antigens on oligodendrocytes and H-2 and Ia antigens on astrocytes by IFN- γ . *a*, Oligodendrocytes viewed with: 1, phase contrast; 2, rhodamine fluorescence (GC⁺) and 3, fluorescein fluorescence (H-2⁺). *b*, Astrocytes viewed with: 1, phase contrast; 2, fluorescein fluorescence (GFAP⁺); 3, rhodamine fluorescence (H-2⁺). *c*, Astrocytes viewed with: 1, phase contrast; 2, fluorescein fluorescence (GFAP⁺); 3, rhodamine fluorescence (Ia⁺).

Methods: Brain cell populations from 2-day-old CBA mice were grown for 6–7 days on coverslips and then exposed to IFN- γ (2 units per ml) for 24 h. The oligodendrocytes were first stained with biotin-conjugated monoclonal antibody against H-2 antigens (11-4-1)^{5,6} and fluorescein-conjugated avidin, followed by monoclonal antibody against galactocerebroside (GC)²¹ and rhodamine-conjugated rabbit anti-mouse immunoglobulin. The astrocytes were stained first with biotin-conjugated monoclonal antibodies against either H-2 (11-4-1)⁵ or Ia (14-4-4S)⁵ antigens, followed by rhodamine-conjugated avidin. The staining was specific in that fluorochrome-conjugated avidin binds only to biotin-conjugated antibodies whereas fluorochrome-conjugated rabbit anti-mouse immunoglobulin did not bind to the monoclonal antibody 11-4-1 which has been conjugated to biotin. The failure of rabbit anti-mouse immunoglobulin to bind biotin-conjugated mouse monoclonal antibodies is of some technical interest and will be reported elsewhere (G.H.W.W. *et al.*, in preparation). All cells were examined by fluorescence microscopy using $\times 1,000$ magnification.

labelling studies showed directly that the percentage of GFAP⁺, GC⁺, Fc⁺ and A2B5⁺ cells that expressed H-2 antigens at the levels detectable by fluorescence microscopy were 100, 100, 85 and 9, respectively. Examples of double-labelled H-2⁺ astrocytes and oligodendrocytes are shown in Fig. 2.

IFN- α/β produced by Semliki Forest virus-infected L cells⁴ was also examined in the same system and was found to be less effective. A concentration of 200 units per ml of IFN- α/β was required to induce detectable levels of H-2 antigens on only 50% of the viable cells (not shown). Comparison of the antiviral titres showed that the H-2-inducing activity of IFN- α/β was at least 200-fold less potent than IFN- γ .

To determine whether IFN- γ had similar effects on brain cells *in vivo*, IFN- γ was injected directly into the brains of 2-day-old mice. After 36 h, cell suspensions were prepared from the brains as before and assayed immediately for expression of H-2 antigens. Cell populations from IFN- γ -injected brains expressed high levels of H-2 antigens (at least 30-fold increase over control) on about 50% of cells (Fig. 3d), whereas <2% of cells isolated from the brains injected with control supernatants expressed detectable levels of H-2 antigens (Fig. 3a). H-2 antigen induction was detectable at 24 h and was maximal 36 h after injection. A maximal effect with a single injection required 25 units per brain. In contrast to the experiments with cultured cells (Fig. 1), treatment of brain with IFN- γ *in vivo* did not

result in the expression of H-2 antigens on all cells. This difference probably reflects the uneven distribution of the IFN- γ injected locally, because the percentage of H-2-positive cells rose to above 80% in mice given two injections into different sites over a 3-day period.

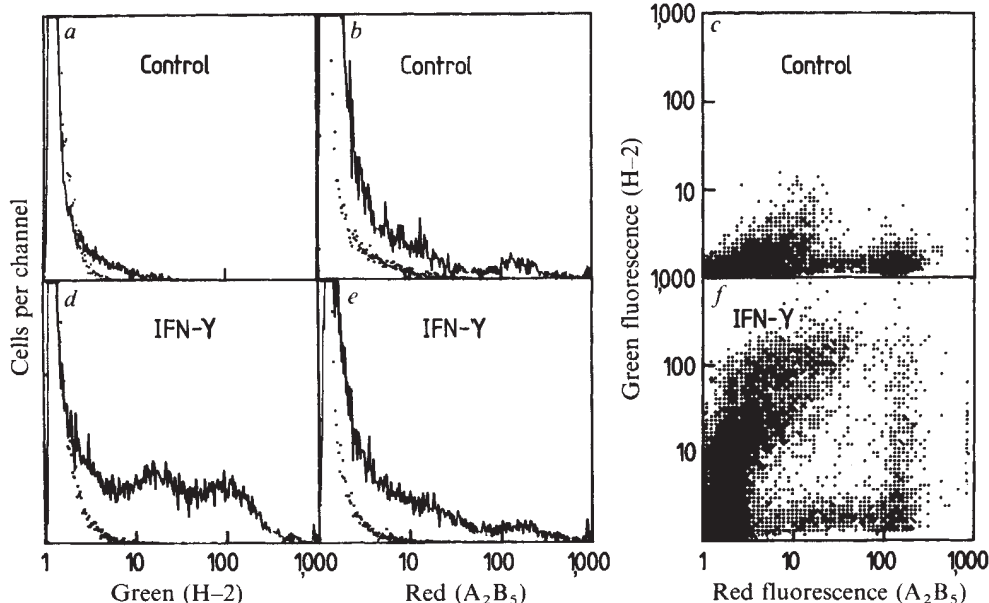
Because neurones could not be readily maintained *in vitro* (Table 1), it was of particular interest to examine the induction of H-2 antigens on neurones *in vivo*. As shown in Fig. 3b, e, 35 $\pm$ 2% of the viable cells freshly isolated from the brains with or without IFN- γ treatment were A2B5⁺. Double-labelling studies showed that 15 $\pm$ 1% of these A2B5⁺ cells from IFN- γ -treated brains expressed detectable levels of H-2 antigens (Fig. 3f) whereas <1% of A2B5⁺ cells from the control brains were H-2⁺ (Fig. 3c). About 30% of the brain cells were NF⁺ and 10 $\pm$ 3% of NF⁺ cells from IFN- γ -treated brains were H-2⁺ by fluorescence microscopy. Double-labelling of H-2⁺ cells with either A2B5 or antibody to neurofilaments is shown in Fig. 4. Other double-labelling experiments showed that the percentages of GFAP⁺, Fc⁺ and GC⁺ cells that expressed detectable levels of H-2 antigens by fluorescence microscopy were 62, 45 and 55, respectively.

We also examined the effect of IFN- γ on the expression of Ia antigens. Freshly isolated brain cells or cells that had been cultured for 24 h with or without IFN- γ did not express detectable levels of Ia antigens. However, when populations of brain

Fig. 3 Induction of H-2 antigens on brain cells *in vivo*. The results show the levels of fluorescein fluorescence indicating the levels of H-2 antigens on cells obtained from the control (a) or IFN- γ -injected (d) brains, and levels of rhodamine fluorescence, indicating the levels of A2B5 staining on cells derived from the control (b) or IFN- γ -injected (e) brains, in each case plotted against cell number. Also shown are the same data from control (c) or IFN- γ -injected brains (f) in which the levels of green (H-2) are plotted against red (A2B5) fluorescence.

Methods: Cerebra and cerebella of 2-day-old BALB/c mice were injected with a total of 50 μ l per brain of control supernatant lacking IFN- γ activity or IFN- γ (50 units per brain). Single-cell suspensions (with 85% viability) were prepared as described for Fig. 1 and were first stained with biotin-conjugated monoclonal antibody against H-2K^dD^d antigens (34-1-2S)^{5,6}

followed by fluorescein-conjugated avidin and then with a monoclonal antibody (A2B5)²³ against the GQ ganglioside, followed by rhodamine-conjugated rabbit anti-mouse immunoglobulin. Propidium iodide was added to stain non-viable cells and allow their exclusion from analysis using the FACS II, and ~25,000 stained cells per sample were analysed. The percentages of the viable cells within the set gates were 85% and 89% of the control and IFN- γ -treated cells, respectively.



cells that had been cultured for 6–7 days and consisted of ~90% astrocytes were treated with IFN- γ (2 units per ml) for a further 24 h, about 10% of the cells expressed Ia antigens as detected by fluorescence microscopy. Double-labelling studies with anti-GFAP and antibodies against Ia antigens showed that the labelled cells were astrocytes; an example is shown in Fig. 2c. In contrast, IFN- α/β (200 units per ml) did not induce detectable Ia antigens on these cells. In cells freshly derived from brains directly injected 36 h previously with IFN- γ , about 5% of cells displayed Ia antigens, compared with less than 1% of cells from the control brains. Double-labelling experiments identified these cells as astrocytes. The significance of these findings is emphasized by the recent finding of Fontana *et al.*²⁶ that astrocytes which displayed these antigens following contact with T cells, could present myelin basic protein to encephalitogenic T-cell lines.

Overall, these results indicate clearly that IFN- γ induces H-2 antigens on astrocytes, oligodendrocytes, microglia and at least some neurones *in vitro* and *in vivo*. Although H-2 antigens could be induced on 99% of the cells that survived 24 h *in vitro*, we could not exclude the existence of subpopulations of cells such as neurones which might not be inducible but were selectively lost during the culture period. Similarly, because of uncertainties about the distribution of IFN- γ *in vivo* it cannot be concluded that negative cells are necessarily uninducible. The experiments both *in vitro* and *in vivo* rely on a similar technique for isolating the brain cells and bias in the recovery of cell types could occur.

T cells have been shown to be actively involved in certain brain lesions^{27–30}, and the pathology of certain virus infections in the brain appears to be caused by virus-specific cytotoxic T cells³¹. Given the absence of H-2 antigens in the normal brain, it might be expected that virus-infected brain cells would escape

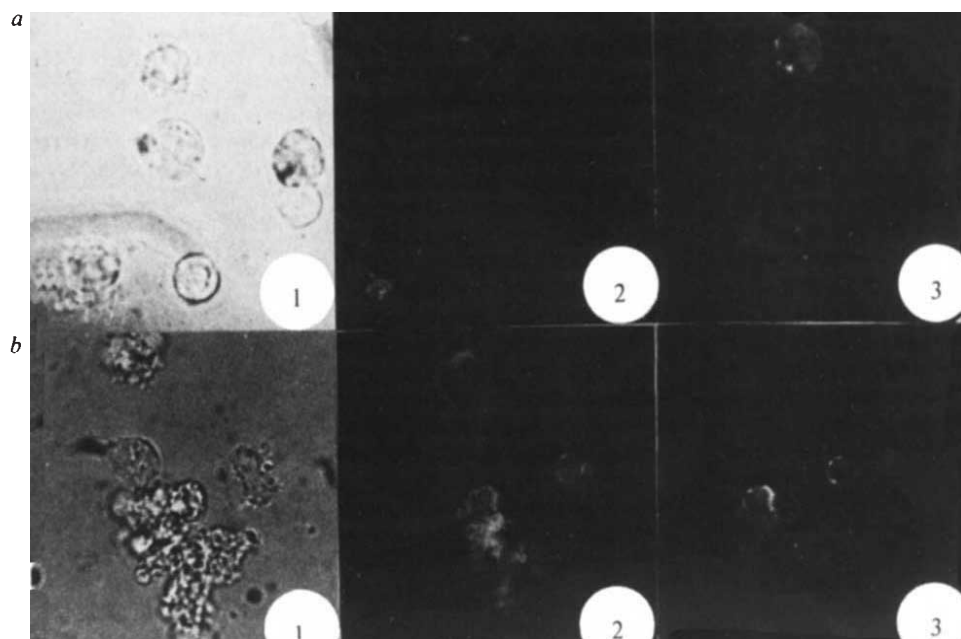


Fig. 4 Induction of H-2 antigens on neurones. Cells from the brains of BALB/c mice injected 36 h previously with IFN- γ (50 units per brain) were prepared and stained with anti-H-2 (34-1-2S) and A2B5 as described in Fig. 3 legend. Another group of H-2⁺ cells was cytocentrifuged, fixed and further stained with antibody which binds to neurofilaments²⁵ as described in Table 1 legend. a, Cells viewed (760-fold magnification) with: 1, phase contrast; 2, rhodamine fluorescence (A2B5⁺); 3, fluorescein fluorescence (H-2⁺). b, Cells viewed ($\times 400$ magnification) with: 1, phase contrast; 2, fluorescein fluorescence (NF⁺); 3, rhodamine fluorescence (H-2⁺).

Table 1 Composition of the brain cell population before and after culture

	Before culture (%)	After culture (%)	
		Without IFN- γ	With IFN- γ
GFAP ⁺ (astrocytes)	37.3 $\pm$ 2.1	63.8 $\pm$ 11.5	61.6 $\pm$ 10.2
GC ⁺ (oligodendrocytes)	1.7 $\pm$ 0.4	1.5 $\pm$ 0.4	1.2 $\pm$ 0.3
Fc ⁺ (microglia)	4.7 $\pm$ 1.1	9.5 $\pm$ 1.8	6.9 $\pm$ 1.1
A2B5 ⁺ (neurones)	38.3 $\pm$ 9.8	8.1 $\pm$ 0.5	8.4 $\pm$ 1.4
NF ⁺ (neurones)	32.3 $\pm$ 5.5	3.7 $\pm$ 4.0	3.3 $\pm$ 3.1

Brain cells, either freshly prepared or cultured for 24 h as described in Fig. 1 legend, were stained for cell type-specific markers and fluorochrome-conjugated anti-mouse or anti-rabbit immunoglobulin. Monoclonal antibodies to the GQ ganglioside (A2B5)²³, galactocerebroside^{20,21} (anti-GC, from Dr Barbara Ranscht)²¹ and Fc receptor (2.4G2)²² (from A. Lopez) were used to identify neurones²³, oligodendrocytes^{20,21} and microglial cells¹⁹, respectively. Astrocytes and neurones were identified by staining with an antiserum against cytoplasmic glial fibrillary acidic protein (GFAP)¹⁹ (rabbit anti-human GFAP; Dako-Santa Barbara) and rabbit antiserum to α -MSH which binds to neurofilaments²³ (Immunonuclear), respectively, while the controls were incubated with normal rabbit serum. For identification of cytoplasmic markers, cells were cytocentrifuged and fixed in 5% acetic acid in ethanol before labelling with rabbit anti-GFAP or rabbit anti-NF serum and fluorescein-conjugated goat anti-rabbit immunoglobulin. All cells were fixed on coverslips, mounted in glycerol, sealed with nail varnish and examined by fluorescence microscopy using 1,000-fold magnification. Dead cells were identified by uptake of propidium iodide and excluded from the analysis. Results are expressed as the mean $\pm$ s.d. from three separate experiments. About 20% of viable cells were not identified; they could be leptomeningial cells, fibroblasts or cells belonging to the indicated lineages but not expressing any of these markers.

recognition and destruction by cytotoxic T cells which can recognize viral antigens only in association with the H-2 antigens of the cell³². However, our results suggest that in the presence of activated T cells or IFN- γ , H-2 antigens can be induced on brain cells, thus rendering them susceptible to lysis by cytotoxic T cells.

We thank Drs P. Gray and D. Goeddel for the gift of IFN- γ , Dr J. L. McKimm-Breschkin for IFN assays, K. Wycherley and S. Liapis for preparation of cells, M. Cozens for help with flow cytometry, and Dr A. W. Harris and Professor G. J. V. Nossal for comments on this paper. This work was supported by the National Health and MRC, the Bushell Trust, the Windermere Hospital Foundation (Australia), National Multiple Sclerosis Society grant RG1435-A-1, and NIH grant AI03958 to G. J. V. Nossal.

Received 27 October 1983; accepted 5 June 1984.

- Williams, K., Hart, D., Fabre, J. & Morris, P. *Transplantation* **29**, 274-279 (1980).
- Vitetta, E. S. & Capra, J. D. *Adv. Immun.* **26**, 147-193 (1978).
- Klein, J., Juretic, A., Baxevanis, C. N. & Nagy, Z. A. *Nature* **291**, 455-460 (1981).
- Wong, G. H. W., Clark-Lewis, I., McKimm-Breschkin, J. L. & Schrader, J. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6989-6993 (1982).
- Wong, G. H. W., Clark-Lewis, I., McKimm-Breschkin, J. L., Harris, A. W. & Schrader, J. W. *J. Immun.* **131**, 788-793 (1983).
- Wong, G. H. W., Clark-Lewis, I., Harris, A. W. & Schrader, J. W. *Eur. J. Immun.* **14**, 52-56 (1984).
- Koch, N., Wong, G. H. W. & Schrader, J. W. *J. Immun.* **132**, 1361-1369 (1984).
- Stee, P. S., Moore, R. N., Johnson, H. M. & Oppenheim, J. J. *J. exp. Med.* **156**, 1780-1793 (1982).
- Kim, K. J., Chaouat, G., Leiserson, W. M., King, J. & De Maeyer, E. *Cell. Immun.* **76**, 253-267 (1983).
- King, D. P. & Jones, P. P. *J. Immun.* **131**, 315-318 (1983).
- Wallach, D., Fellous, M. & Revel, M. *Nature* **299**, 833-836 (1982).
- Pober, J. S. *et al. J. exp. Med.* **157**, 1339-1353 (1983).
- Basham, T. Y. & Merigan, T. C. *J. Immun.* **130**, 1492-1494 (1983).
- Kelley, V. E., Fiers, W. & Strom, T. B. *J. Immun.* **132**, 240-245 (1984).
- Wong, G. H. W., Clark-Lewis, I., Hamilton, J. A. & Schrader, J. W. *J. Immun.* (in the press).
- Wong, G. H. W. & Schrader, J. W. *Lymphokines* **11** (in the press).
- Wong, G. H. W., Bartlett, P. F., Clark-Lewis, I. & Schrader, J. W. *J. Neuroimmun.* (in the press).

- Gray, P. W. & Goeddel, D. V. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5842-5846 (1983).
- Raff, M. C. *et al. Brain Res.* **174**, 283-308 (1979).
- Raff, M. C., Mishez, R. & Field, R. L. *Nature* **274**, 813-816 (1978).
- Ranscht, B., Clapham, P. A., Price, J., Noble, M. & Seifert, W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2709-2713 (1982).
- Unkeless, J. C. *J. exp. Med.* **150**, 580-596 (1979).
- Eisenbarth, G. S., Walsh, F. S. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4913-4917 (1979).
- Raff, M. C., Miller, R. H. & Noble, M. *Nature* **303**, 390-396 (1983).
- Drager, U. C., Edwards, D. L. & Kleinschmidt, J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6408-6412 (1983).
- Fontana, A., Fierz, W. & Wekerle, H. *Nature* **307**, 273-276 (1984).
- Traugott, U., Reinherz, E. L. & Raine, C. S. *Science* **219**, 308-310 (1983).
- Traugott, U., Reinherz, E. L. & Raine, C. S. *J. Neuroimmun.* **4**, 201-221 (1983).
- Nyland, H., Matre, R., Mørk, S., Bjerke, J. R. & Naess, A. *New Engl. J. Med.* **307**, 1643-1644 (1982).
- Brinkman, C. J. J., Ter Laak, H. J., Hommes, O. R., Poppema, S. & Delmotte, P. *New Engl. J. Med.* **307**, 1644-1645 (1982).
- Cole, G. A., Nathanson, N. & Prendergast, R. A. *Nature* **238**, 335-337 (1972).
- Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701-702 (1974).
- McDevitt, H. O., Delovitch, T. L., Press, J. L. & Murphy, D. B. *Transplant. Rev.* **30**, 197-235 (1979).
- Kirby, W. N. & Parr, E. L. *J. Histochem. Cytochem.* **27**, 746-750 (1979).
- Parr, E. L. *J. exp. Med.* **150**, 1-9 (1979).

Protein kinase C activation of physiological processes in human neutrophils at vanishingly small cytosolic Ca²⁺ levels

F. Di Virgilio*, D. P. Lew† & T. Pozzan*

* CNR Unit for the Study of Physiology of Mitochondria, c/o Institute of General Pathology, Via Loredan 16, 35100 Padova, Italy
† Infectious Diseases Unit, Department of Internal Medicine, University of Geneva, CH-1211 Geneva 4, Switzerland

It has long been assumed that a rise in cytosolic free Ca²⁺, [Ca²⁺]_i, is a necessary and sufficient event for the stimulation of a variety of cellular processes^{1,2}. The development of a technique which allows monitoring of [Ca²⁺]_i in small intact cells³ has led to a critical revision of this simple postulate⁴⁻⁷. We have recently shown that in neutrophils, Ca²⁺-ionophore-induced elevations of [Ca²⁺]_i, quantitatively similar to those caused by chemotactic peptides, are ineffective in stimulating cell responses⁸, which suggests that an additional signal is required for receptor-mediated activation. Here we show that subthreshold concentrations of phorbol myristate acetate (PMA) and of a Ca²⁺ ionophore can quantitatively mimic the effect of a physiological agonist. However, PMA at higher concentrations can trigger NADPH-oxidase activity, exocytosis and protein phosphorylation, even when [Ca²⁺]_i is lowered 10-20 times below the normal resting level. These results strongly suggest that activation of protein kinase C is sufficient, by itself, to induce NADPH-oxidase activation and exocytosis of secondary granules in neutrophils.

Recently, Rink *et al.*⁵ have shown that in platelets, the effects of thrombin on cell function can be mimicked by a combination of subthreshold levels of a Ca²⁺ ionophore and of PMA or 1-oleoyl-, 2-acetyl-glycerol (OAG). Both compounds are known to substitute for diacylglycerol, the supposed physiological activator of the ubiquitous phospholipid-dependent Ca²⁺-activated protein kinase C (ref. 9). A similar synergistic effect between PMA and the Ca²⁺ ionophore, ionomycin, can be obtained in human neutrophils (Fig. 1, trace b). Addition of ionomycin in Ca²⁺-free medium, which is not stimulatory *per se* (trace c), induces a classical respiratory burst in neutrophils pretreated with a subthreshold concentration of PMA. Trace a of Fig. 1 shows that in the same batch of quin2-loaded cells, PMA slightly decreases [Ca²⁺]_i, while ionomycin causes an abrupt and transient rise of [Ca²⁺]_i. In Ca²⁺-supplemented or Ca²⁺-free medium, a synergistic effect occurred between subthreshold levels of ionomycin and PMA with regard to the secretion of specific granules (vitamin B₁₂ binding protein) or, in the presence of cytochalasin B, of primary granules (β -

glucuronidase) (not shown; see also ref. 10). Two important differences were noted between the synergistic effects of PMA and ionomycin on O_2 consumption and on secretion. Secretion starts immediately after ionomycin addition and is complete within about 1 min, whereas O_2 consumption starts about 30 s after ionomycin addition, that is, when $[Ca^{2+}]_i$ is falling towards basal levels, and survives the return of $[Ca^{2+}]_i$ to resting level.

The most popular interpretation¹¹ of the synergism between Ca^{2+} ionophores and protein kinase C activators is that PMA or OAG activates protein kinase C which increases the Ca^{2+} sensitivity of processes triggered by the Ca^{2+} ions themselves. An alternative possibility, however, is that protein kinase C, once activated by OAG or PMA, is sufficient *per se* to trigger some cellular metabolic responses¹². In this case, the role of Ca^{2+} will be indirect, synergizing with subthreshold levels of PMA or OAG to activate protein kinase C and/or sensitizing some other intracellular process to subthreshold activation of protein kinase C (for example membrane fusion). One way to distinguish between these two possibilities would be to decrease $[Ca^{2+}]_i$ to very low levels (10^{-8} – 10^{-9} M) at which all known Ca^{2+} -dependent processes are practically inactive (see, for example, refs 13–18) and to see whether a response can still be elicited with PMA. One problem with this approach is that *in vitro* protein kinase C is itself considered to be Ca^{2+} dependent even in the presence of PMA¹⁹. The simplest way of detecting the activity of protein kinase C in intact cells is to look at protein phosphorylation and, indeed, a change in the pattern of protein phosphorylation is among the first detectable effects of PMA on intact neutrophils²⁰. Very low $[Ca^{2+}]_i$ can be obtained by loading cells with quin2 in Ca^{2+} -free medium containing EGTA (ref. 3), conditions under which Ca^{2+} homeostasis cannot be maintained. In neutrophils, quin2 loading in Ca^{2+} -free medium plus EGTA results in a decrease of $[Ca^{2+}]_i$ to 5–10 nM, that is, 10–30 times below resting values (see also Fig. 2). No significant difference in the pattern of protein phosphorylation induced by PMA at resting $[Ca^{2+}]_i$ (100 nM) or at low $[Ca^{2+}]_i$ (8 nM) was observed (data not shown). Thus, even at $[Ca^{2+}]_i$ as low as 8 nM, intense protein phosphorylation can be obtained with PMA, presumably a result of protein kinase C activation. The question then arises whether the cellular responses that can be activated by protein kinase C can still occur at negligible $[Ca^{2+}]_i$.

Figure 2 (lower trace) shows that in neutrophils loaded with quin2 in Ca^{2+} -free medium plus EGTA, $[Ca^{2+}]_i$ 5×10^{-9} M, addition of 20 nM PMA potently stimulates O_2 consumption. Ionomycin does not cause any further stimulation of O_2 consumption or release of Ca^{2+} , showing that the depletion of intracellular Ca^{2+} stores was complete. When Ca^{2+} is added to the cell suspension, its cytosolic concentration rises to micromolar levels but O_2 consumption is unaffected. Furthermore, while the dependence of O_2 consumption rate on PMA concentration is affected by $[Ca^{2+}]_i$, maximal stimulation is practically independent of it (Fig. 3). This result suggests that $[Ca^{2+}]_i$ influences the activity of NADPH oxidase indirectly, through

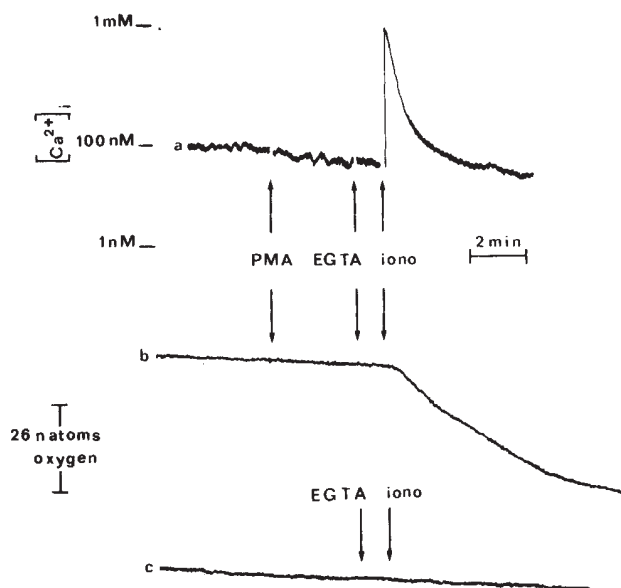


Fig. 1 Ionomycin-induced release of Ca^{2+} from intracellular stores triggers NADPH-oxidase activation in neutrophils pretreated with subthreshold PMA concentrations. Trace a, $[Ca^{2+}]_i$ changes; traces b, c, O_2 consumption in two aliquots of the same batch of quin2-loaded human neutrophils. The suspending medium contained 140 mM NaCl, 5 mM KCl, 1 mM $MgSO_4$, 0.5 mM $CaCl_2$, 1 mM NaH_2PO_4 , 5.5 mM glucose, 20 mM HEPES pH 7.4 (37 °C). 1 nM PMA, 5 mM EGTA and 0.2 μ M ionomycin were added where indicated. Cell number, 2.5×10^6 ml⁻¹. Quin2 content, 0.15 nmol per 10^6 cells. Human neutrophils were obtained from healthy donors and isolated (>95% pure) according to standard procedures⁸. O_2 consumption and fluorescence measurements were performed with an oxygen Clark electrode and a Perkin Elmer 650-40 fluorimeter, respectively. Both the fluorimeter cuvette holder and the oxygen electrode glass vessel were thermostated (37 °C) and magnetically stirred. Excitation and emission wavelengths were 399 ± 3 and 492 ± 10 nm, respectively. To minimize light scattering artefacts, two cutoff filters UV D 25 and UV 35 for excitation and emission, respectively, were used. Quin2 loading and calibration of fluorescence as a function of $[Ca^{2+}]_i$ were performed as described previously^{3,8}. The rise in $[Ca^{2+}]_i$ induced by ionomycin in this experiment was much larger than previously reported in the same conditions⁸. This was obtained by reducing the cellular quin2 content, thus minimizing the Ca^{2+} buffering by the indicator.

protein kinase C, rather than by affecting the enzyme itself. This finding is consistent with the observation that maximal O_2^- production by inside-out vesicles isolated from preactivated alveolar macrophages can occur at 10^{-9} M Ca^{2+} (ref. 21).

Phorbol esters not only trigger the production of toxic O_2 metabolites in phagocytic cells but also have pleiotropic effects on cell functions. In particular, they can cause secretion of exocytotic granules in this and other cell types^{22,23}. Exocytosis involves fusion between membranes and rearrangement of cytoskeletal components and it has been generally assumed that such processes require Ca^{2+} . It was therefore of great interest to see whether PMA could also induce exocytosis at $[Ca^{2+}]_i$ levels of $\leq 10^{-8}$ M. Table 1 shows that the release of vitamin B₁₂ binding protein is stimulated by PMA even in cells in which cytosolic Ca^{2+} has been buffered at 10^{-8} M. The secretion is very little affected by cytochalasin B and, surprisingly, is slightly larger than in the controls with a normal resting $[Ca^{2+}]_i$. Similarly, PMA-induced secretion of β -glucuronidase, in the presence of cytochalasin B, is not affected by reducing $[Ca^{2+}]_i$ to 10 nM and is hardly above 10% of the total content after 15 min, both at resting and at low $[Ca^{2+}]_i$ (not shown). However, when $[Ca^{2+}]_i$ is elevated above resting level with a dose of ionophore which is not stimulatory *per se*, the release of β -glucuronidase induced by PMA reaches values as high as 50%

Table 1 Release of vitamin B₁₂ binding protein

Stimulus	10^{-8} M $[Ca^{2+}]_i$		10^{-7} M $[Ca^{2+}]_i$	
	+CB	-CB	+CB	-CB
PMA	37 $\pm$ 0.7	35 $\pm$ 1.0	36 $\pm$ 2.8	22 $\pm$ 2.5
FMLP	2 $\pm$ 0.3	4 $\pm$ 1.5	21 $\pm$ 2.5	12 $\pm$ 2.5

CB, cytochalasin B. Quin2 content, 0.69 nmol per 10^6 cells. Cell number, 2.5×10^6 per ml. PMA and FMLP concentrations were 30 nM and 1 μ M, respectively. Cytochalasin B was 5 μ g ml⁻¹. The values are expressed as % of the total content after 15 min of incubation with either stimulus subtracted from basal release. Values are means $\pm$ s.d. of triplicates of a typical experiment. The loading was performed as described for Fig. 2. The cells were stimulated in media of the same composition as used for the loading. 10^{-7} M is the level of $[Ca^{2+}]_i$ before addition of the stimulant. $[Ca^{2+}]_i$ decreased to 9×10^{-8} M or increased to 6×10^{-7} M after PMA or FMLP, respectively. At 10^{-8} M $[Ca^{2+}]_i$, neither PMA nor FMLP affected $[Ca^{2+}]_i$.

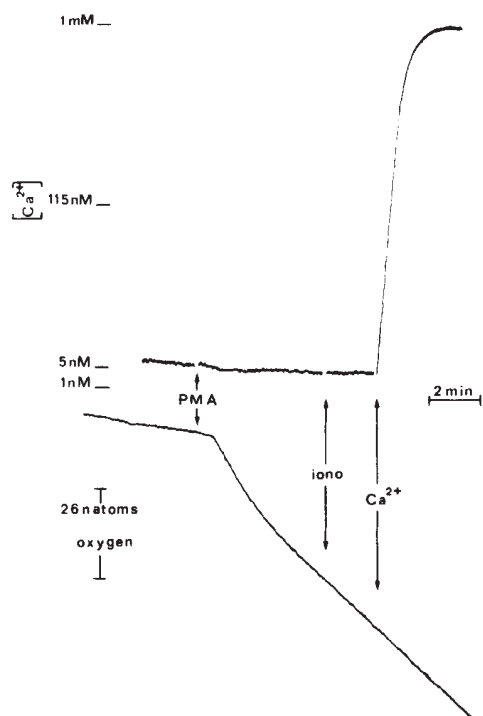


Fig. 2 Activation of oxygen consumption by PMA at low $[Ca^{2+}]_i$. Upper trace, $[Ca^{2+}]_i$ changes; lower trace, O_2 consumption. The suspending medium was as for Fig. 1 except that $CaCl_2$ was omitted and EGTA was present. The loading was performed as usual, but in a medium without Ca^{2+} , containing 1 mM EGTA, that is, $[Ca^{2+}]_0 = 10^{-9}$ M. 20 nM PMA, 0.2 μ M ionomycin and 2 mM $CaCl_2$ were added where indicated. Cell number, 2.5×10^6 ml $^{-1}$. Quin2 content, 0.55 nmol per 10^6 cells. The small decrease of fluorescence on addition of 20 nM PMA, but not of 1 nM (see Fig. 1), is due to an autofluorescence artefact. As 5 nM $[Ca^{2+}]_i$, fMet-Leu-Phe neither stimulates O_2 consumption nor increases $[Ca^{2+}]_i$ (not shown).

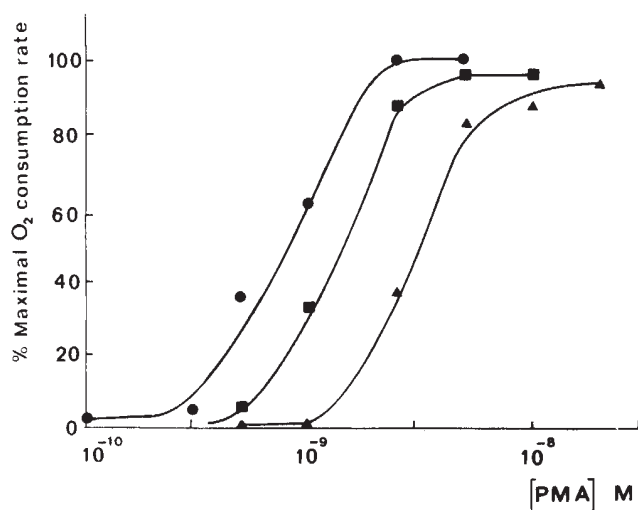


Fig. 3 Dependence of O_2 consumption on the PMA concentration at three different $[Ca^{2+}]_i$. ●, Cells loaded with quin2 in Ca^{2+} medium, kept in the presence of Ca^{2+} throughout the experiment and treated with 0.2 μ M ionomycin 3 min before the addition of PMA ($[Ca^{2+}]_i = 800$ nM). ■, Cells loaded with quin2 in Ca^{2+} medium and stimulated with PMA in the same medium ($[Ca^{2+}]_i = 100$ nM). ▲, Cells loaded with quin2 in Ca^{2+} -free medium + 1 mM EGTA ($[Ca^{2+}]_i = 10$ nM). Cell number, 2.5×10^6 ml $^{-1}$. Quin2 content, 0.55 nmol per 10^6 cells. In this experiment, in contrast to Fig. 1, PMA 1 nM slightly stimulated O_2 consumption, probably due to minor differences among the various cell batches.

(see also ref. 10). We stress that the difference in release of primary and secondary granules induced by PMA is only quantitative and that massive release of primary granules is a laboratory artefact, because it occurs only in the presence of cytochalasin B.

In contrast, secretion induced by the chemotactic peptide fMet-Leu-Phe is practically abolished when $[Ca^{2+}]_i$ is reduced to 10 nM. One might speculate that at low $[Ca^{2+}]_i$, fMet-Leu-Phe induces the formation of only part (diacylglycerol?) of the effective signal which is rapidly inactivated. Only when there is a simultaneous elevation of $[Ca^{2+}]_i$ is the metabolic response fully expressed. On the other hand PMA, which is only slowly metabolized¹⁹, leads to a practically irreversible cell activation and does not require Ca^{2+} .

Two main conclusions can be drawn from these results: (1) PMA-induced protein phosphorylation can occur independently of $[Ca^{2+}]_i$; (2) PMA-dependent activation of O_2 consumption and secretion may be modulated by cytosolic Ca^{2+} concentration, although it is possible to stimulate these responses by increasing PMA concentration at any $[Ca^{2+}]_i$.

The identification of the intracellular target of PMA as protein kinase C in many cellular systems^{24,25}, the stimulatory activity of OAG^{5,10}, the modulation of PMA- or OAG-dependent activation by $[Ca^{2+}]_i$ (ref. 5) and the recent report by Helfman *et al.*²⁶ of a Ca^{2+} -phospholipid-dependent protein kinase activity in the membrane fraction of neutrophils, point to the involvement of protein kinase C in the responses we have investigated. The main question raised by our results is whether reactions occurring at 10^{-8} M Ca^{2+} or less can be considered Ca^{2+} independent. We know of no report of a Ca^{2+} -dependent reaction with a K_m for Ca^{2+} lower than 50–100 nM. We, however, observe maximum activation of exocytosis and NADPH oxidase at 5–10 nM $[Ca^{2+}]_i$. The ability of PMA to stimulate these metabolic responses at less than 10 nM $[Ca^{2+}]_i$ suggests that protein phosphorylation by protein kinase C does not simply increase the Ca^{2+} sensitivity of Ca^{2+} -dependent processes but itself triggers responses until now considered strictly Ca^{2+} dependent.

We thank Mr G. Ronconi for technical assistance, Mrs F. Mazzari for typing, Dr C. Montecucco and Dr G. B. Wollheim for helpful discussions. Ionomycin was a kind gift from C. M. Liu of Hoffmann La Roche. This work was supported by a grant from the Italian Ministry of Public Education (40%), by a CNR grant (progetto finalizzato oncologia) and by grant 3.986-082 from the Swiss NSF. D.P.L. was a recipient of a Max Cloetta Career Development Award.

Received 10 April; accepted 14 June 1984.

- Rubin, R. P. *Calcium and Cellular Secretion* (Plenum, New York, 1982).
- Gomperts, B. D. in *Biological Membranes* Vol. 5 (ed. Chapman, D.) 284–340 (Academic, New York, 1984).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *J. Cell Biol.* **94**, 325–334 (1982).
- Rink, T. J., Smith, S. W. & Tsien, R. Y. *FEBS Lett.* **148**, 21–26 (1982).
- Rink, T. J., Sanchez, A. & Hallam, I. J. *Nature* **305**, 317–319 (1983).
- Pozzan, T., Arslan, A., Tsien, R. Y. & Rink, T. J. *J. Cell Biol.* **94**, 335–340 (1982).
- Meldolesi, J., Huttner, W. B., Tsien, R. Y. & Pozzan, T. *Proc. natn. Acad. Sci. U.S.A.* **81**, 620–624 (1984).
- Pozzan, T., Lew, D. P., Wollheim, C. B. & Tsien, R. Y. *Science* **221**, 1413–1415 (1983).
- Nishizuka, Y. *Trends biochem. Sci.* **8**, 13–16 (1981).
- Kajikawa, N. *et al. Biochem. biophys. Res. Commun.* **116**, 743–750 (1983).
- Kaibuchi, K. *et al. J. biol. Chem.* **258**, 6701–6704 (1983).
- Michell, R. *Trends biochem. Sci.* **8**, 263–265 (1983).
- Yin, H. L. & Stossel, T. P. *J. biol. Chem.* **255**, 9490–9493 (1980).
- Klee, C. B., Crouch, T. H. & Krinks, M. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6270–6273 (1979).
- Pershadsingh, H. A., Landt, M. & McDonald, J. M. *J. biol. Chem.* **255**, 8983–8986 (1980).
- Niggli, V., Adunyah, E. S. & Carafoli, E. *J. biol. Chem.* **256**, 8588–8592 (1981).
- Gillis, J. M., Thomason, D., Lefevre, J. & Kretsinger, R. H. *J. Muscle Res. Cell Motility* **3**, 377–398 (1982).
- Burger, D., Stein, E. A. & Cox, J. A. *J. biol. Chem.* **258**, 14733–14739 (1983).
- Castagna, M. *et al. J. biol. Chem.* **257**, 7847–7851 (1982).
- Schneider, C., Zanetti, M. & Romeo, D. *FEBS Lett.* **127**, 4–8 (1981).
- Lew, D. P. & Stossel, T. P. *J. clin. Invest.* **67**, 1–9 (1981).
- Estensen, R. D., White, J. G. & Holmes, B. *Nature* **248**, 347–348 (1974).
- Gunther, G. R. *J. biol. Chem.* **256**, 12040–12045 (1981).
- Niedel, J. E., Kuhn, L. J. & Vandenbark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **8**, 263–265 (1983).
- Sando, J. J. *et al. Fedn Proc.* **42**, 2250–2253 (1983).
- Helfman, D. M., Appelbaum, B. D., Vogler, W. R. & Kuo, J. F. *Biochem. biophys. Res. Commun.* **111**, 847–853 (1983).

Stylites, a vascular land plant without stomata absorbs CO₂ via its roots

Jon E. Keeley*, C. Barry Osmond† & John A. Raven‡

* Department of Biology, Occidental College, Los Angeles, California 90041, USA

† Desert Research Institute, University of Nevada, Reno, Nevada 89506, USA

‡ Department of Biological Sciences, University of Dundee, Dundee DD1 4HN, UK

Photosynthetic organs of most higher plants normally have access to atmospheric CO₂ through stomatal pores which also serve as variable valves to control the loss of H₂O vapour which accompanies CO₂ uptake¹. The acquisition of stomata is commonly thought to have been a crucial development permitting 'conquest' of land and direct access of plants to atmospheric CO₂. Only in desert stem succulents during drought do stomata remain so tightly closed in the light that the photosynthetic tissues are dependent on internal CO₂ generated through the photosynthetic pathway known as crassulacean acid metabolism². Functional stomata are absent in submerged aquatic plants and in non-vascular land plants (for example, mosses) which are normally covered by a water film. Although it is now clearly established that some aquatic plants assimilate large amounts of CO₂ from the sediment via roots³⁻⁵, terrestrial plants are thought to assimilate only insignificant amounts of CO₂ via this path⁶. Here we report on a terrestrial plant, *Stylites andicola*, which lacks stomata and is unable to exchange gas with the aerial atmosphere. Rather, it derives nearly all of its photosynthetic carbon through its roots. In addition, this species possesses characteristics of crassulacean acid metabolism.

S. andicola (Isoetaceae) is a rare fern ally found in small populations in the high Andes of Peru⁷. We studied a small population in the vicinity of Lago de Junin (4,135 m) in December 1982. There, as well as at other sites⁸, it forms dense colonies of hundreds of plants most commonly on hummocky ground elevated above the water level of the surrounding (seasonal) bog. The substrate is a highly decomposed peat (largely derived from *Stylites*) exceeding 1 m depth and overlying limestone. Interstitial groundwater pH was 6.4–6.6 and alkalinity ranged from 50 to 70 g m⁻³ CaCO₃.

Stylites produces an evergreen rosette (2–3 cm across) of leaves (2–4 cm long) arising from an elongated corm which may be branched (Fig. 1). These plants are embedded in the peat so

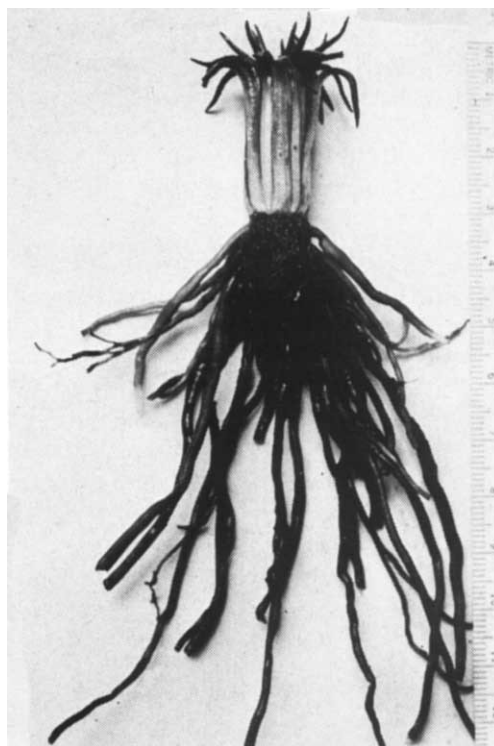


Fig. 1 *Stylites andicola*; scale is a 12-cm ruler.

that the bulk of the plant is underground; the only exposed surfaces are the upper-third chlorophyllous portion of the leaves which constitute a small part of the total biomass. Average plant dry weight = 0.829 g (s.d. = 0.19, $n = 4$); green portions of leaves = 6.6%, white portions = 18.0%, stems = 44.0% and roots = 31.4%.

Stylites leaves lack stomata and are covered with a thick continuous cuticle which is essentially impermeable to CO₂ and H₂O vapour. Measurements with a dual isotope porometer⁹ (on plants maintained as described in Table 1) gave CO₂ conductances (mm s⁻¹) in the morning and afternoon of 0.001 (s.d. = 0.001, $n = 4$) and 0.005 (s.d. = 0.007, $n = 4$), respectively. These conductances are nearly an order of magnitude lower than those observed for desert cacti with closed stomata².

When the photosynthetic tissues were exposed to ¹⁴CO₂ in the light, negligible rates of ¹⁴C incorporation were observed (Table 1). However, when the root system was exposed to ¹⁴CO₂, up to 50 times the amount of ¹⁴C was incorporated in the photosynthetic tissues. In field conditions this difference could be even greater as, during transplanting, the original root systems died and were only partially recovered at the time of these experiments. The root/green tissue ratio for oven-dried plants from the field was 5.5 ± 3.3 ($n = 9$), in contrast to a ratio of 2.5 ± 1.1 ($n = 12$) for the plants in Table 1.

In natural conditions the chlorophyll-containing tissues of *Stylites* showed a nighttime increase and a daytime decrease in titratable acidity and malic acid content (Table 2), characteristic of crassulacean acid metabolism (CAM) which is common in aquatic species in the Isoetaceae¹⁰. However, the amount of CO₂ recycled through CAM each night was only equivalent to ~1 h of photosynthesis when roots were exposed to ¹⁴CO₂. Carbon isotope discrimination, often helpful in identifying photosynthetic pathways and CO₂ sources in CAM, does not help to evaluate these processes in this case. The $\delta^{13}\text{C}$ for green leaves (–25.7‰) was similar to that of the peat (–26.6‰) and groundwater carbonate (–25.0‰). These data are consistent with the assimilation of peat-derived CO₂ in a closed system, either through CAM or C₃ pathways of photosynthesis. The chlorophyllous tips of *Stylites* leaves contain sufficient activities of ribulosebiphosphate carboxylase (70–134 $\mu\text{mol per mg}$

Table 1 ¹⁴CO₂ uptake by chlorophyllous portions of *Stylites* leaves

		CO ₂ uptake by green leaves* ($\mu\text{mol CO}_2$ per mg chlorophyll h ⁻¹)	
		¹⁴ CO ₂ fed to leaves	¹⁴ CO ₂ fed to roots
Light	(Moist)	0.74 ± 0.14 (3)	39.75 ± 6.14 (3)
	(Wet)	1.12 ± 1.05 (2)	18.78 ± 0.79 (2)
Dark	(Moist)	0.26 ± 0.16 (2)	0.51 ± 0.07 (2)

Leaves and roots were isolated in separate chambers similar to those used in aquatic plant studies^{3,4}. In all trials green portions of leaves were kept in air and roots in 25 mM NaH₂PO₄, 2 mM NaHCO₃ buffer at pH 6.5, which approximated field carbon and pH conditions. NaH¹⁴CO₃ (4.7 mCi μmol^{-1}) was injected into either the root medium or a small tray of HCl in the leaf chamber to produce a final activity of 2 $\mu\text{Ci ml}^{-1}$. Plants were incubated for 2 h at 25 °C with 1,000 $\mu\text{einsteins m}^{-2} \text{s}^{-1}$. Gentle stirring was applied in both the top and bottom of the chamber. Tissues were killed in boiling 80% ethanol, ground in a glass homogenizer, dried and resuspended in distilled H₂O and counted. Values are the mean ± s.d. (n = number of experiments). Before experiments, the plants had been maintained in either moist but drained peat or water-saturated peat at 13 °C during the day and 7 °C during the night on a 12-h photoperiod.

* These tissues contained 1.0 ± 0.2 mg chlorophyll per g fresh weight.

Table 2 Titratable acidity (to pH 6.4) and malic acid content in green portions of leaves of *Stylites* collected from moist sites and wet or temporarily submerged conditions in December at the Junin site

Site	Acidity (μ equiv. per g fresh weight)		Malic acid content (μ mol per g fresh weight)	
	a.m.	p.m.	a.m.	p.m.
Moist	20 $\pm$ 10	9 $\pm$ 1	34 $\pm$ 6	17 $\pm$ 1
Wet	32 $\pm$ 4	10 $\pm$ 2	41 $\pm$ 7	24 $\pm$ 5

For methods, see ref. 10 ($n = 2$).

chlorophyll h^{-1}) to support the measured rates of $^{14}CO_2$ fixation (Table 1) and phosphoenolpyruvate carboxylase activity (24–34 μ mol per mg chlorophyll h^{-1}) was adequate to support observed rates of acid accumulation.

The path of carbon movement from roots to leaf tips in *Stylites* is unknown. The leaves are connected to the roots with a vascular trace⁸ but nothing is known of the gas exchange characteristics of these tissues. The leaves and roots have abundant air spaces; the former have four large air canals and more than 25% of the cross-sectional area is air space. Chloroplasts are especially abundant in cells bordering these air spaces. Due to the collapse of cortical tissue early in development, the roots are relatively hollow; 75% of the cross-sectional area is air space. The youngest roots are initiated at the top of the corm and extend through the corm to the middle of the rosette of leaves. However, older roots may be separated from leaves by a centimetre or more of spongy corm tissue.

This study raises intriguing questions concerning the photosynthetic physiology, as well as ecological and phylogenetic relationships of these primitive land plants. In the natural habitat, colonies of *Stylites* are interspersed amongst a dense array of other plants. Some degree of convergence is evident in the superficially similar rosette arrangement of photosynthetic parts of *Stylites* and associated species. However, in the major aspects discussed here, *Stylites* is unique; all the other rosette-forming species associated with *Stylites* possess stomata and finely divided roots with a dense cortex.

Whether the morphology and photosynthetic physiology of *Stylites* represents an important link between aquatic plants and primitive land plants, or whether it is merely an interesting variant, merits further careful assessment. Although cuticles of leaf-like structures which lack stomata are known from the Devonian/Silurian, it is not always easy to find stomata even in well preserved fossils¹¹. Organisms such as *Stylites* may represent relics of many marginally successful compromises from which present day terrestrial photosynthetic systems were selected. With their principal functional counterpart, the seasonally sealed stem succulents, *Stylites* share an ability to carry out CAM. Their capacity for CAM supports the notion that this pathway of photosynthetic carbon assimilation has long been part of the metabolic background of higher plants¹². Because *Stylites* lacks stomata and exhibits CAM, this unusual organism does not conform to the hypothesis that CAM may have evolved from stomatal guard cell metabolism¹³.

This research was supported by grant DEB-8206887 (J.E.K.), and by J. Teeri (University of Chicago). L. Bates, S. Keeley, C. Lopez, C. Osmond, I. Ting and D. Walker contributed to this project.

Received 24 April; accepted 12 June 1984.

- Farquhar, G. D. & Sharkey, T. D. *A. Rev. Pl. Physiol.* **33**, 317–345 (1982).
- Kluge, M. & Ting, I. P. *Crassulacean Acid Metabolism* (Springer, Berlin, 1978).
- Wium-Andersen, S. *Physiologia Pl.* **25**, 245–248 (1971).
- Sondergaard, M. & Sand-Jensen, K. *Aquat. Bot.* **6**, 1–12 (1979).
- Richardson, K., Griffiths, H., Reed, M. L., Raven, J. A. & Griffiths, N. M. *Oecologia* **61**, 115–121 (1983).
- Berquist, N. O. *Bot. Notiser* **117**, 249–258 (1964).
- Amstutz, E. *Ann. Mo. bot. Gdn* **44**, 121–123 (1957).
- Rauh, W. & Falk, H. *Sber. heidelb. Akad. Wiss.* **1**, 1–83; **2**, 1–160 (1959).
- Johnson, H. B., Rowlands, P. G. & Ting, I. P. *Photosynthetica* **13**, 409–418 (1979).
- Keeley, J. E. *Am J. Bot.* **69**, 254–257 (1982).
- Edwards, E., Edwards, D. S. & Rayner, R. In *The Plant Cuticle* (eds Cutler, D. F., Alvins, K. L. & Price, C. E.) 341–361 (Academic, New York, 1982).
- Osmond, C. B. in *Proc. VI int. Photosynth. Congr.* Vol III, 557–564 (Junk, The Hague, 1984).
- Cockburn, W. *Pl. Cell Envir.* **4**, 417–418 (1981).

Associative cellulolysis and dinitrogen fixation by co-cultures of *Trichoderma harzianum* and *Clostridium butyricum*

D. A. Veal

Agricultural and Food Research Council Letcombe Laboratory, Wantage, Oxon OX12 9JT, UK

J. M. Lynch*

Glasshouse Crops Research Institute, Worthing Road, Littlehampton, West Sussex BN17 6LP, UK

Lignocelluloses are a major source of carbon into ecosystems and represent the principal constituent of agricultural, forestry and municipal wastes. The large proportion of cellulose and hemicellulose in these wastes leads to a high C:N ratio which often results in N limitation during decomposition. This could be overcome if the decomposer organisms combined both the cellulolytic (cellulase) and N_2 -fixing (nitrogenase) functions. With the exception of bacteria isolated from the specialized environment of marine shipworms^{1,2}, no organism possessing both cellulase and nitrogenase has yet been isolated from nature or genetically engineered. We now report the cooperative degradation of cellulose in which the fungus *Trichoderma harzianum* provides the cellulase function and an obligately anaerobic bacterium, *Clostridium butyricum*, provides the nitrogenase function. These co-cultures utilized cellulose as the sole carbon source for N_2 fixation, resulting in a substantial increase in the rate of substrate decomposition compared with the fungus alone. The co-cultures developed in apparently aerobic environments, demonstrating that aerobes and anaerobes can co-exist, the aerobe presumably providing respiratory protection to the anaerobe.

Complex materials such as plant residues require a variety of enzymes for their complete degradation which in natural environments is usually provided by mixed microbial communities. Communities with both cellulase and nitrogenase functions are often associated with the degradation of cellulosic residues which have a high C:N ratio. N_2 fixation has been found associated with the decomposition of straw^{3–6}, detritus from salt marsh plants⁷ and wood^{8,9}. In these residues, restricted aeration seems to be a prerequisite for substantial N_2 fixation whereas the total exclusion of O_2 inhibits N_2 fixation^{6,10}. Isolations from a straw-decomposing N_2 -fixing system⁶ reveal that aerobic fungi represent the major cellulase producers and anaerobic bacteria the major N_2 fixers. Restricted aeration in these systems is often achieved by waterlogging¹¹, N_2 fixation being maximal at the interface between the aerobic and anaerobic environments^{6,12}.

Using co-cultures of the cellulolytic fungus *T. harzianum* and the anaerobic N_2 -fixer *C. butyricum*, N gains, measured by elemental analysis and confirmed by acetylene reduction assays, of 18.5 μ g ml^{-1} , equivalent to 7.87 mg N fixed per g carbon lost, were achieved using cellulose as the sole carbon source (Table 1). Neither acetylene reduction nor N gains to the system were found in pure cultures of *T. harzianum* or *C. butyricum*.

The amount of substrate decomposed and the formation of cell protein were much greater in the co-culture than with fungus growing alone. It seems likely that the increased decomposition rate by the co-cultures is a consequence of the bacterium using the end products of fungal cellulolysis, relieving any end product inhibition of fungal cellulases. The fixation of N also helps to alleviate N deficiency, allowing greater microbial development, as demonstrated by the increase in cell protein concentration. The N_2 fixed in the mixed culture (18.5 μ g $N\ ml^{-1}$) is equivalent to lowering the C:N ratio to 71.9. Degradation of the substrate by pure cultures of *T. harzianum* causes the C:N ratio to fall

* To whom correspondence should be addressed.

Table 1 The decomposition of cellulose by *Trichoderma harzianum* in pure culture and in co-culture with *Clostridium butyricum*

Inoculant	Measured N gain ($\mu\text{g ml}^{-1}$)	C ₂ H ₂ reduced ($\mu\text{mol ml}^{-1}$)	Weight loss (mg ml^{-1})	Final C:N	Cell protein ($\mu\text{g ml}^{-1}$)	Acetate ($\mu\text{mol ml}^{-1}$)	Butyrate ($\mu\text{mol ml}^{-1}$)	Final pH
a <i>Clostridium butyricum</i>	0	0	0	107.8	0	0	0	7.0
b <i>Trichoderma harzianum</i>	0	0	3.40	76.0	106	0	0	6.4
c <i>Trichoderma harzianum</i> + <i>Clostridium butyricum</i>	18.5	1.958	6.43	29.6	235	8.14	9.03	4.5
			$P \leq 0.05$ b, c	0.001 a-c	0.001 b, c			0.001 a-c

Organisms were grown in 100 ml (pH 7) mineral salts medium¹⁷ contained in a conical flask (capacity 325 ml). $(\text{NH}_4)_2\text{SO}_4$ (0.1 g l^{-1}) and yeast extract (0.1 g l^{-1}) were added as an initial N source and as a source of growth factors. Total carbon was determined using a Carlo Erba elemental analyser model 1106. The initial C:N ratio of the complete medium was 231.8 before addition of the N source and 107.8 after its addition. Wet ball-milled Whatman CF11 cellulose powder was the carbon source. Oxoid Bacteriological Agar (2 g l^{-1}) was added to prevent convectional movements in the medium and to keep cellulose in suspension in the static cultures. A spore inoculum of the fungus was inoculated into the medium and shaken at 200 r.p.m. for 7 days before inoculation of the bacterium; subsequently, the flasks were not shaken. Acetylene reduction¹⁸ was measured every 48 h and integrated over the 20-day culture period. Acetate and butyrate were determined by gas chromatography¹⁹ and cell protein was determined colorimetrically using Coomassie brilliant blue²⁰ with bovine serum albumin as standard. The cultures were freeze-dried to determine weight loss and the N content of the residue measured using an elemental analyser⁵. N gain was determined from the difference in total N content between the culture and uninoculated controls.

from 107.8 to 76.0 solely as a result of C loss. The fall in C:N of the mixed culture is far greater, from 107.8 to 29.6. This fall is a consequence of the increase in the total N content as a result of N₂ fixation and the stimulation in substrate decomposition (C loss). Using ¹⁵N-labelled cells of *C. butyricum*, it has been demonstrated that bacterial N is available to *T. harzianum*. Thus, N₂ fixed by *C. butyricum* could be available for cellulase production and thereby stimulate substrate decomposition.

Not only were rates of decomposition increased by co-cultures, but also the products which result from cellulolysis were different, as only in the mixed cultures were acetate, butyrate and vigorous gas production observed.

Consistent with the findings of Jensen¹³ and Vartiavaara¹⁴, we found that the addition of a small amount of N (0.1 mg ml^{-1} $(\text{NH}_4)_2\text{SO}_4$) to the medium was required in order to establish good fungal growth before inoculation with the N₂-fixer. Without this added N, fungal development was poor and con-

sequently growth and N₂ fixation by the bacterium were low. The added N does not suppress N₂ fixation because it is converted into fungal biomass, removing its inhibitory effect on nitrogenase.

These mixed cultures were grown in static conditions in aerobic environments. The fungus permitted the growth of the anaerobe in these conditions presumably by intercepting available O₂. Thus, not only does the fungus make the substrate available to the N₂ fixer, it also protects the anaerobe from O₂.

In a second series of experiments, *T. harzianum* was replaced by an exogenous supply of the enzyme cellulase. In the absence of O₂, growth and N₂ fixation took place, as indicated by butyrate production and acetylene reduction (Fig. 1). At enzyme concentrations less than 2×10^{-4} enzyme units (EU) ml^{-1} , growth and nitrogenase activity were dependent on the cellulase concentration and were probably limited by the rate of cellulolysis. In the total absence of cellulase, *C. butyricum* was unable to utilize cellulose and neither cellulase nor yeast extract in the absence of cellulose support substantial growth or acetylene reduction. Thus, *C. butyricum* is dependent on the hydrolysis of cellulose by cellulase to produce the simple sugars (cellobiose and glucose) which it is capable of fermenting.

In nature, N₂ fixation is expensive in energetic (ATP) terms and in free-living bacteria is generally considered to be limited by available carbohydrate¹⁵. The present observations demonstrate the potential of cellulosic compounds to satisfy that need. Indeed, we have already provided preliminary evidence that associative N₂ fixation can occur with straw as a substrate⁵ and have more recently isolated a range of organisms capable of forming such associations⁶.

Our ultimate objective is to exploit cellulolytic N₂-fixing associations to produce novel composts for use in agriculture and horticulture¹⁶. The use of such associations offers an alternative to the genetic manipulation of single species.

We thank S. H. T. Harper and S. J. Chapman for helpful comments. The work was supported in part by EEC contract RUW-033-UK. D.A.V. is an AFRC research student.

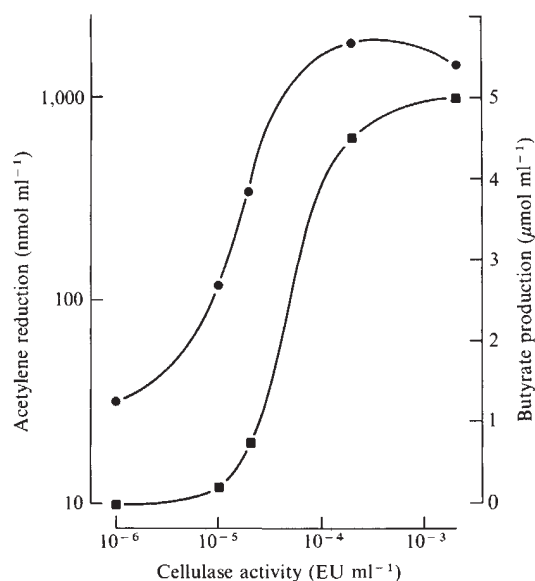


Fig. 1 Acetylene reduction (●) and butyrate production (■) by *C. butyricum* and supplied with cellulase (BDH; ex *Trichoderma viride*) growing on a medium introducing cellulose as the sole carbon source. No activity occurred in the absence of cellulase. Cultures were contained in static glass tubes (capacity 18 ml) containing 5 ml of sloppy cellulose agar as described in Table 1 legend except that no $(\text{NH}_4)_2\text{SO}_4$ was added.

Received 13 April; accepted 11 June 1984.

1. Carpenter, E. J. & Culliney, J. I. *Science* **187**, 551-552 (1975).
2. Waterbury, J. B., Calloway, C. B. & Turner R. D. *Science* **221**, 1401-1403 (1983).
3. Yoneyama, T., Lee, K. K. & Yoshida, T. *Soil Sci. Pl. Nutr.* **23**, 287-295 (1977).
4. Rice, W. A. & Paul, E. A. *Can. J. Microbiol.* **18**, 715-723 (1972).
5. Lynch, J. M. & Harper, S. H. T. *J. gen. Microbiol.* **129**, 251-253 (1983).
6. Harper, S. H. T. & Lynch, J. M. *J. appl. Bact.* (in the press).
7. Haines, E. B. & Hanson, R. B. *J. exp. mar. Biol. Ecol.* **40**, 27-40 (1979).
8. Aho, P. E., Seidler, R. S., Evans, H. J. & Nelson, A. D. in *Proc. Int. Symp. Nitrogen Fixation* (eds Newton, W. E. & Nynman, C. J.) 629-690 (Washington State University Press, Pullman, 1976).

9. Silvester, B., Sollins, T., Verhoeven, T. & Cline, S. P. *Can. J. For Res.* **12**, 646–652 (1982).
10. Barrow, N. J. & Jenkinson, D. S. *Pl. Soil* **16**, 258–262 (1962).
11. Buresh, R. J., Casselman, M. E. & Patrick, W. H. *Adv. Agron.* **33**, 149–192 (1980).
12. Magdoff, F. R. & Bouldin, D. R. *Pl. Soil* **33**, 49–61 (1970).
13. Jensen, H. L. *Proc. Linn. Soc. N.S.W.* **66**, 239–249 (1941).
14. Vartiavaara, U. *J. Sci. agric. Soc. Finl.* **10**, 241–264 (1938).
15. Postgate, J. R. & Hill, S. in *Microbial Ecology: A Conceptual Approach* (eds Lynch, J. M. & Poole, N. J.) 191–213 (Blackwell, Oxford, 1979).
16. Lynch, J. M., Harper, S. H. T., Chapman, S. J. & Veal, D. A. in *Microbial Methods for Environmental Biotechnology* (eds Grainger, J. M. & Lynch, J. M. (Academic, New York, in the press).
17. Dalton, H. & Postgate, J. R. *J. gen. Microbiol.* **56**, 307–319 (1969).
18. Postgate, J. R. in *Methods in Microbiology* Vol. 6B in (eds Norris, J. R. & Ribbons, D. W.) 343–356 (Academic, London, 1972).
19. Lynch, J. M. *J. appl. Bact.* **42**, 81–87 (1977).
20. Sedmak, J. J. & Grossberg, S. E. *Analyt. Biochem.* **79**, 544–552 (1977).

c-myc transcript is induced in rat liver at a very early stage of regeneration or by cycloheximide treatment

Reiko Makino, Kenshi Hayashi & Takashi Sugimura

Biochemistry Division, National Cancer Center Research Institute, 1-1 Tsukiji 5 chome, Chuo-ku, Tokyo 104, Japan

In rats, partial hepatectomy induces reasonably synchronized DNA replication in the remaining liver after approximately 20 h¹. Events occurring during the earlier stages of liver regeneration are of interest because they may tell us how cells *in vivo* respond when they move from a differentiated resting state (G₀ phase) to a proliferative state. We report here that the expression of the c-myc oncogene is increased up to 10–15-fold of the normal level within 1–3 h after partial hepatectomy. This expression begins to decrease rapidly after 4 h and has returned to less than double the normal level after 8 h, at which time replicative DNA synthesis has still not begun. A still larger increase in c-myc transcription (~600-fold) is observed in the liver when protein synthesis is inhibited by an injection of cycloheximide. These findings suggest the existence of a short-lived protein that is synthesized soon after partial hepatectomy, and which suppresses the expression of c-myc.

Rats were killed at various times after partial hepatectomy² and total RNA was extracted from their regenerating liver. Figure 1 shows the results of Northern blot analysis³ of the RNA with the 3'-half fragment of v-myc⁴ as a probe. Essentially similar results were obtained when a cloned fragment containing exon 1 of the rat c-myc gene (manuscript in preparation) was

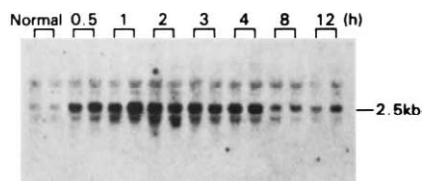


Fig. 1 Levels of c-myc transcripts in the liver at various times after partial hepatectomy. Male Sprague-Dawley rats (200–250 g) were partially hepatectomized under ether anaesthesia following the method of Higgins and Anderson². The rats were killed at the indicated times after the operation and total RNA was extracted from the liver by the guanidinium thiocyanate-hot phenol method¹⁸. The RNA (10 µg per lane) was separated on 1.2% agarose gel containing 6% HCHO₃, blotted onto a nitrocellulose filter and hybridized to a nick-translated 3'-half SalI-PstI fragment of v-myc⁴. Each lane represents an RNA sample from one animal. After exposure, the probe was removed and the RNA was rehybridized⁷ to ³²P-labelled v-ras^H and mouse rDNA sequences. These probes were insert fragments of plasmids BS9 (ref. 19) and p6.6 (ref. 20), respectively. The size of the c-myc transcript was estimated from the position of its band relative to that of rRNA.

used as a probe. As shown in Fig. 1, the amount of the 2.5-kilobase transcript of c-myc had already increased 30 min after the operation; it reached a maximum at 1–3 h and then decreased. The bands were traced with a densitometer and quantitated from their peak areas. Figure 2 compares these values with those of normal liver, together with the time course of incorporation of ³H-thymidine into the acid-insoluble fraction of the liver, when injected intraperitoneally 1 h before death. At 30 min, the amount of c-myc transcript was 10-fold that in control liver, increased to a maximum of 10–15-fold at 1 h and decreased after 4 h. The decrease was rapid and c-myc transcripts had returned to approximately normal levels after 8 h, although DNA synthesis had not begun to increase.

Fausto and Shank⁵ reported that c-myc expression was increased three- to five-fold at 12–18 h after partial hepatectomy. In our results over the same period, levels of c-myc expression were also high, although they were never more than double that in normal liver. Examination of c-myc expression at earlier stages of the regeneration revealed a conspicuous peak soon after partial hepatectomy, whereas the previous report did not determine c-myc expression at that time. We also examined the expression of the Harvey ras gene (c-ras^H) during liver regeneration using the same filters as for analysis of c-myc expression. In accordance with earlier reports^{5–7}, we observed an increase in c-ras^H expression, but this became evident after 8 h, and peaked at a level two to three times that in normal liver, at about 30 h after partial hepatectomy (data not shown).

In vitro stimulation of B lymphocytes, T lymphocytes or cultured fibroblasts with their respective specific mitogens is known to induce an immediate increase in the expression of c-myc⁸. This increase is temporary, and the expression returns to the uninduced level by the time DNA replication starts. It

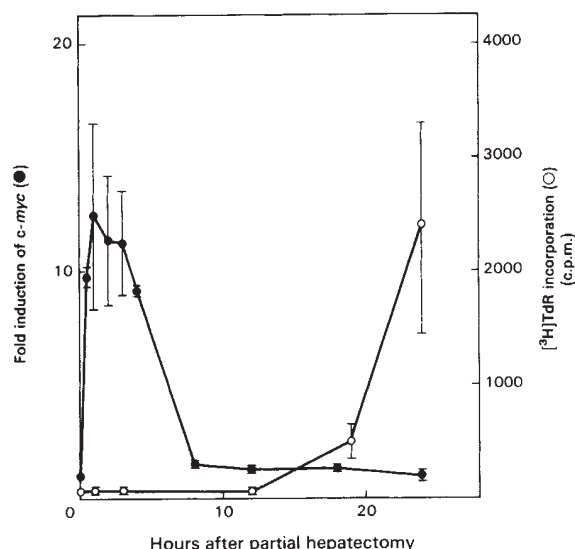


Fig. 2 Relationship between the appearance of c-myc transcript and DNA synthesis in regenerating liver. The X-ray film shown in Fig. 1 and the autoradiogram of rRNA of the same filter were scanned with a densitometer. The amount of c-myc transcript in each sample relative to that of 28S rRNA of the same lane was determined and expressed relative to the value for control liver. The RNA to DNA ratio did not change appreciably (data not shown) during the first 19 h of liver regeneration. Because rRNA represents the majority of the total RNA, the relative amount of c-myc mRNA per DNA and therefore per cell in each sample, can be approximated by the relative ratio of c-myc transcript to 28S rRNA. DNA synthesis was measured as follows. Rats were injected intraperitoneally with ³H-thymidine (TdR) at 100 µCi per rat 1 h before death. Their livers were homogenized in 20-fold excess of the PK buffer described by Favaloro *et al.*²¹. 100 µl of the homogenates were spotted onto a glass fibre filter and their acid-insoluble radioactivities were counted. The points shown are the average of duplicates with the ranges indicated.

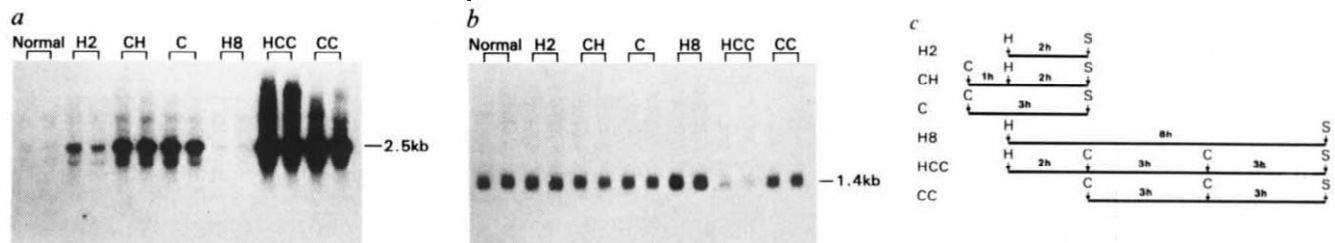


Fig. 3 Effects of cycloheximide on *c-myc* transcription. Rats were either treated with cycloheximide²² (50 mg per kg, intraperitoneally) or partially hepatectomized, or both, according to the protocols shown in *c*. Then *c-myc* (*a*) or *c-ras*^H (*b*) transcripts in total liver RNA were analysed as described in Fig. 1 legend. *c*, Protocol of treatments. Arrows labelled C, H and S denote times of cycloheximide treatment, partial hepatectomy and death, respectively. The relative level of *c-myc* transcript was quantitated on X-ray films of various exposure times. The levels in HCC and CC were found to be 60 times that in H2 (killed 2 h after partial hepatectomy), which was elevated to 10 times that in normal liver. Therefore, in HCC and CC, *c-myc* transcription was increased 600-fold over the level in normal liver.

has also been shown that *c-myc* induction is not blocked by inhibition of protein synthesis but is instead enhanced in mitogen-stimulated cells *in vitro*.

The present findings show that the same is true for an *in vivo* system in which differentiated resting cells are stimulated to proliferate. We examined the effect of inhibition of protein synthesis on the expression of *c-myc* in regenerating liver under the protocol indicated in Fig. 3c. As shown in Fig. 3a, we observed a 100-fold increase in *c-myc* transcription in samples from rats treated with cycloheximide 1 h before partial hepatectomy and killed 2 h after the operation (CH). The amount of *c-ras*^H transcript was not significantly changed by cycloheximide treatment (Fig. 3b). Therefore, the increase in the amount of *c-myc* transcript observed in cycloheximide-treated liver is likely to be the result of enhanced synthesis rather than of stabilization of mRNA in general⁹. However, other possibilities such as specific stabilization of *c-myc* mRNA cannot be excluded. Interestingly, a similar increase has also been found in a sample from a rat without partial hepatectomy treated 3 h previously with cycloheximide (C). This effect of cycloheximide alone was prolonged and was even increased by about 600-fold 6 h after treatment with and without partial hepatectomy (HCC and CC, see Fig. 3 legend for calculation). Thus, the mode of induction of *c-myc* by cycloheximide seems to be different from that in regenerating liver, where the induction is temporary and the extent of induction is increased up to 10–15-fold. Enhanced induction by cycloheximide was also observed when treatment was preceded by partial hepatectomy (HCC). Hence, inhibition of protein synthesis seems to block the switch-off of *c-myc* transcription observed at 4 h or later after partial hepatectomy.

As shown here and in an earlier report⁸, *c-myc* transcription is increased immediately after the cells have been stimulated to proliferate, but this expression of *c-myc* stops soon after it has reached a maximum. Because inhibition of protein synthesis enhances the induction of *c-myc*, it seems likely that *c-myc* is repressed by a short-lived protein that becomes abundant soon after the onset of the proliferative process.

Based on observations in Burkitt's lymphoma cells (in which *c-myc* genes were translocated to immunoglobulin gene loci), Leder *et al.*¹⁰ have elaborated several possible models concerning the regulatory mechanisms of the *c-myc* gene. Among these, the autoregulatory model is the simplest, yet it is consistent with the facts that *c-myc* is expressed transiently following inductive stimulation and that its mRNA is induced by the inhibition of protein synthesis. In *Escherichia coli*, analogous autoregulatory mechanisms have been described for a stress protein, *dnaK*, and for a repressor protein in the SOS function, *lexA*, which becomes extensively but transiently expressed after inductive stimuli^{11,12}. The induction following the proliferative signal can be explained by supposing that the signal activates a process, such as modification of *c-myc* protein, thereby abolishing its repressor activity. Expression of *c-myc* at a low but distinct level in various tissues¹³

is also consistent with autoregulation. Another possibility is that *c-myc* expression is repressed by some other 'early gene' products¹⁴. Such gene products are expressed transiently at an early stage in mitogen-stimulated fibroblasts and the levels of their mRNA are enhanced by inhibition of protein synthesis.

We have found previously that in all chemically induced primary hepatomas examined, the level of *c-myc* transcript was three to five times that in normal liver or normal liver tissue adjacent to the tumour⁷. Altered regulation of *c-myc* in some B-cell lymphomas and in other cancer cells as a result of translocation, viral enhancer insertion or gene amplification are well established phenomena^{15,16}. Recently, Campisi *et al.*¹⁷ observed that in non-transformed A31 mouse fibroblast cells, *c-myc* could be induced by growth stimulation, but that it was constitutive in two chemically transformed derivatives of A31. These facts strongly suggest that altered regulation, and perhaps abnormal increase in the expression of *c-myc*, might prevent the cells from entering the G₀ phase and thus lead to their infinite growth. Further studies on the regulation of *c-myc* expression may increase our understanding of the processes involved in the development of cancer.

We thank Drs J. M. Bishop, R. W. Ellis and R. Kominami for supplying the plasmids pv-myc, BS9 and p6.6, respectively, also Dr M. Terada for valuable discussion and encouragement. This work was supported in part by Grants-in-Aid for Cancer Research from the Ministry of Education, Science and Culture and the Ministry of Health and Welfare of Japan, and the Princess Takamatsu Cancer Research Fund.

Received 25 April; accepted 25 June 1984.

1. Bucher, N. L. R. *Int. Rev. Cytol.* **15**, 245–300 (1963).
2. Higgins, G. M. & Anderson, R. M. *Archs Path.* **12**, 186–202 (1931).
3. Goldberg, D. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5794–5798 (1980).
4. Alitalo, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 100–104 (1983).
5. Fausto, N. & Shank, P. R. *Hepatology* **3**, 1016–1023 (1983).
6. Goyette, M., Petropoulos, C. J., Shank, P. R. & Fausto, N. *Science* **219**, 510–512 (1983).
7. Makino, R., Hayashi, K., Sato, S. & Sugimura, T. *Biochem. biophys. Res. Commun.* **119**, 1096–1102 (1984).
8. Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 605–610 (1983).
9. Stiles, C. D., Lee, K. L. & Kenney, F. T. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2634–2638 (1976).
10. Leder, P. *et al. Science* **222**, 765–771 (1983).
11. Tilly, K., McKittrick, N., Zylicz, M. & Georgopoulos, C. *Cell* **34**, 641–646 (1983).
12. Little, J. W. & Mount, D. W. *Cell* **29**, 11–22 (1982).
13. Gonda, T. J., Sheiness, D. K. & Bishop, J. M. *Molec. cell. Biol.* **2**, 617–624 (1982).
14. Cochran, B. H., Reffel, A. C. & Stiles, C. D. *Cell* **33**, 939–947 (1983).
15. Land, H., Parada, L. F. & Weinberg, R. A. *Science* **222**, 771–778 (1983).
16. Hayashi, K., Makino, R. & Sugimura, T. *Gann* **75**, 475–478 (1984).
17. Campisi, J., Gray, H. E., Pardee, A. B., Dean, M. & Sonenshein, G. E. *Cell* **36**, 241–247 (1984).
18. Feramisco, J. R., Smart, J. E., Burridge, K., Helfman, D. M. & Thomas, G. P. *J. biol. Chem.* **257**, 11024–11031 (1982).
19. Ellis, R. W. *et al. J. Virol.* **36**, 408–420 (1980).
20. Mishima, Y., Kominami, R., Honjo, R. & Muramatsu, M. *Biochemistry* **19**, 3780–3786 (1980).
21. Favalaro, J., Treisman, R. & Kamen, R. *Meth. Enzym.* **65**, 718–749 (1980).
22. Chaudhuri, S. & Lieberman, I. *J. biol. Chem.* **243**, 29–33 (1968).

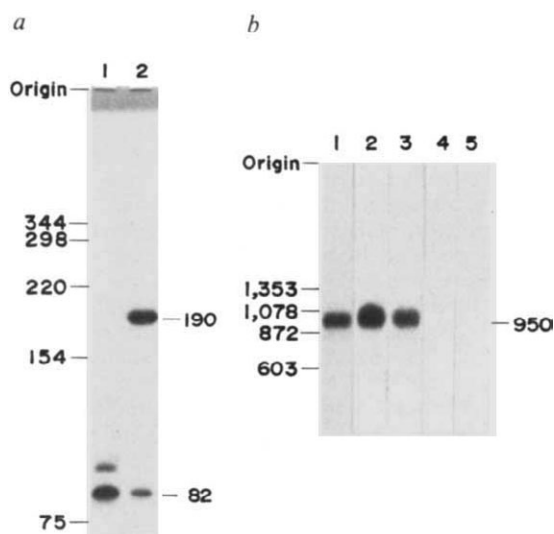


Fig. 2 Identification of the 5' terminus of human preproCDD-ANF mRNA by primer extension (a) and blot-hybridization analysis of human and rat cardiac RNAs (b). An autoradiogram of a is shown: lane 1, primer extension with yeast tRNA; lane 2, primer extension with human atrial poly(A) RNA. The sizes of the primer (82 nucleotides) and the primer-extended cDNA (~190 nucleotides) are shown on the right side of the autoradiogram. Note that the latter size is very similar to that of the cloned cDNA sequence between the 5' end of the primer DNA and the 3' end of the antimesage strand of the cDNA sequence (184 nucleotides, see Fig. 1). An autoradiogram of b is shown: lane 1, human atrial RNA; lane 2, rat left atrial RNA; lane 3, rat right atrial RNA; lane 4, rat left ventricular RNA; lane 5, rat right ventricular RNA. The size markers used in a and b are the 5'-end-labelled *HinfI* cleavage products of pBR322 DNA and the 5'-end-labelled *HaeIII* cleavage products of Φ X174 DNA respectively.

Methods: For a, the ~240-base pair (bp) *RsaI* fragment containing the 5'-sequence of the cDNA insert and its flanking plasmid DNA was isolated and labelled with ^{32}P at the 5' end. The labelled fragment was cleaved by *SacI*, and the *SacI*-*RsaI* fragment (residues 8-85) was isolated and used as a primer; note that the size of the primer strand is 82 nucleotides long (see Fig. 1). The primer DNA was denatured in a solution containing 80% formamide, 40 mM PIPES pH 6.4, 0.4 M NaCl and 1 mM EDTA at 75 °C for 10 min and then hybridized to human atrial poly(A) RNA (5 μg) in the same solution at 42 °C for 3 h. The RNA-DNA hybrid was precipitated by ethanol and subjected to reverse transcriptase reaction³¹. For b, the rat heart was dissected into four parts (the left and the right atrium and the left and the right ventricle) and each part was pooled from 20 rats. Total RNA was extracted²⁵ and 20 μg of rat RNA isolated from each part and 5 μg of human total atrial RNA were denatured with 1 M glyoxal and 50% dimethyl sulphoxide³², electrophoresed on 2.0% agarose gel and transferred to diazophenylthioester-cellulose paper³³. The 581-bp *SacI*-*PstI* fragment (residues 8-588) was used as a probe; the probe was labelled by nick-translation³⁴ with [α - ^{32}P]dCTP.

region of the clone pAF48 indicates that the cloned cDNA sequence covers almost the full length of the mRNA sequence (Fig. 2a). The first 25 amino acid residues, starting with the initiator methionine and preceding the CDD sequence, include a large number of hydrophobic amino acids, including sequences of four and two consecutive leucine residues, and end with an amino acid that has a small neutral side chain, alanine. This structure probably represents a signal peptide similar to that found in other secretory proteins^{16,17}.

Human preproCDD-ANF is schematically illustrated in Fig. 3. Human preproCDD-ANF consists of 151 amino acids, including a putative signal peptide of 25 amino acids, and has a calculated molecular weight of 16,395. The two biologically significant atrial peptides, CDD and ANF, are contained at the N-terminal and C-terminal regions of the prohormone respec-

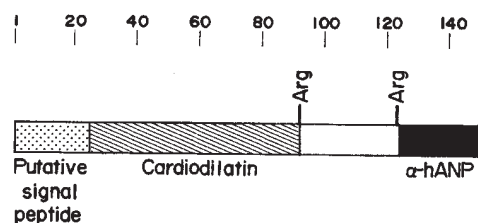


Fig. 3 Schematic representation of the structure of human preproCDD-ANF. The sequences of CDD and α -hANF are indicated by the shaded and solid boxes, respectively, and the putative signal peptide by the stippled box; the C-termini of CDD and the signal peptide are not definite. Amino acid numbers are given above the structure.

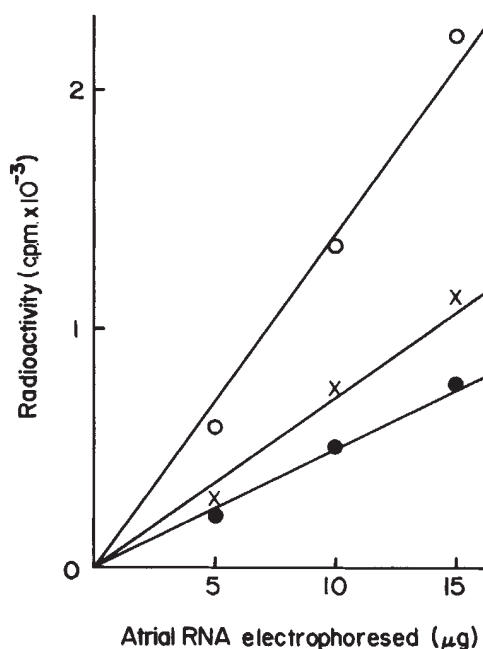


Fig. 4 Effect of water deprivation on the level of rat atrial preproCDD-ANF mRNA. Rats with free access to water (○) and those at 2 days (×) and 4 days (●) after water deprivation.

Methods: Total atrial RNA was extracted from each group of six experimental rats and the indicated amounts of RNAs were analysed by the blot-hybridization technique described in Fig. 2 legend. After autoradiography, the portion corresponding to a hybridization band in the diazophenylthioester-cellulose paper was cut and its radioactivity was counted in a toluene scintillator solution.

tively. A single arginine residue precedes α -hANF; it presumably serves as a proteolytic processing signal, as is observed in some other peptide hormone precursors^{18,19}. Similarly, the arginine residue at position 92 could be used as a proteolytic cleavage site for the formation of human CDD; this would give a CDD peptide with a molecular weight (7,368) similar to that of porcine CDD (7,500, ref. 9).

It has previously been shown that the number of atrial secretory granules containing the ANF peptides varies with changes in water-electrolyte balance²⁰, although neither the site of synthesis of the ANF and CDD peptides nor the regulation of their synthesis is well understood. We therefore investigated the distribution and expression of rat preproCDD-ANF mRNA by the blot-hybridization technique; Fig. 2b shows the autoradiogram of the hybridization analysis. Both human and rat atrial RNA

showed a single hybridization band with a similar mobility, corresponding to a size of 950 nucleotides. This band was observed with RNAs isolated from the left and right atria but not with those from the left and right ventricles. No hybridization-positive band was detected with RNAs obtained from the rat brain, liver or kidney (data not shown). Thus, we conclude that preproCDD-ANF mRNA is selectively synthesized in the atrium. The abundance of preproCDD-ANF mRNA in the atrium is noteworthy: although not quantified precisely, it was estimated to be between 0.5% and 1.0% (based on a comparison of the relative hybridization signal of the RNA with that of a known amount of the cDNA fragments run parallel with the RNA on the electrophoresis gel).

In the experiment shown in Fig. 4, rats were deprived of water for 2 and 4 days and the effect of water deprivation on the level of atrial preproCDD-ANF mRNA was examined by blot-hybridization analysis. The radioactivity measured in the hybridization band showed a linear relationship with the amount of RNA applied to the electrophoresis gel, thus enabling the analysis to be performed quantitatively. Atrial preproCDD-ANF mRNA decreased on water deprivation, and the levels of the mRNA in the rats at 2 and 4 days after water deprivation were about two and three times lower than that in the control rats respectively (Fig. 4). The amount of total atrial RNA extracted was not significantly changed after 2 days' water deprivation but was reduced to about two-thirds of that of the control rats after 4 days' water deprivation. Thus, the level of atrial preproCDD-ANF mRNA is specifically reduced in the early stage of water deprivation, and after further water deprivation, both this level and the total amount of preproCDD-ANF mRNA are markedly lowered.

The results presented here indicate that a major mRNA species in the cardiac atrium encodes the common precursor for CDD and ANF, which may have multi-functional roles in the regulation of extracellular fluid volume and blood pressure. Furthermore, the marked reduction of the amount of preproCDD-ANF mRNA on water deprivation suggests that the water-electrolyte balance plays an important part in the regulation of the preproCDD-ANF gene expression. Further investigations concerning the expression of the preproCDD-ANF gene would thus be fruitful for understanding the molecular basis of not only the physiological maintenance of but also pathological disturbances of water-electrolyte balance and blood pressure.

Since the submission of this manuscript, three groups²¹⁻²³ have published the nucleotide sequences of the cloned cDNAs for rat and human preproCDD-ANF. The nucleotide sequences for human preproCDD-ANF reported by Oikawa *et al.*²³ and by our group agree with each other, except for slight differences in the lengths of the extreme 5' and 3' regions.

We thank Mr Ko-ichi Zenita for technical assistance. The investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Institute of Physical and Chemical Research and the Mitsubishi Foundation.

Received 22 May; accepted 29 June 1984.

- de Bold, A. J., Borenstein, H. B., Veress, A. T. & Sonnenberg, H. *Life Sci.* **28**, 89-94 (1981).
- de Bold, A. J. *Can. J. Physiol. Pharmacol.* **60**, 324-330 (1982).
- Sonnenberg, H., Cupples, W. A., de Bold, A. J. & Veress, A. T. *Ann. N. Y. Acad. Sci.* **372**, 213-214 (1981).
- Briggs, J. P., Steipe, B., Schubert, G. & Schnermann, J. *Pflügers Arch. ges. Physiol.* **395**, 271-276 (1982).
- Pollock, D. M. & Banks, R. O. *Clin. Sci.* **65**, 47-55 (1983).
- Currie, M. G. *et al. Science* **221**, 71-73 (1983).
- de Bold, A. J. *Proc. Soc. exp. Biol. Med.* **170**, 133-138 (1982).
- Trippodo, N. C., Macphree, A. A., Cole, F. E. & Blakesley, H. L. *Proc. Soc. exp. Biol. Med.* **170**, 502-508 (1982).
- Forssmann, W. G. *et al. Anat. Embryol.* **168**, 307-313 (1983).
- Flynn, T. G., de Bold, M. L. & de Bold, A. J. *Biochem. biophys. Res. Commun.* **117**, 859-865 (1983).
- Currie, M. G. *et al. Science* **223**, 67-69 (1984).
- Kangawa, K. & Matsuo, H. *Biochem. biophys. Res. Commun.* **118**, 131-139 (1984).
- Thibault, G. *et al. FEBS Lett.* **167**, 352-356 (1984).
- Misono, K. S., Kukumi, H., Grammer, R. T. & Inagami, T. *Biochem. biophys. Res. Commun.* **119**, 524-529 (1984).
- Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976).

- Blöbel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 852-862 (1975).
- Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N. Y. Acad. Sci.* **343**, 1-16 (1980).
- Land, H., Schütz, G., Schmale, H. & Richter, D. *Nature* **295**, 299-303 (1982).
- Deschenes, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 726-730 (1984).
- de Bold, A. J. *Proc. Soc. exp. Biol. Med.* **161**, 508-511 (1979).
- Yamanaka, M. *et al. Nature* **309**, 719-722 (1984).
- Maki, M. *et al. Nature* **309**, 722-724 (1984).
- Oikawa, S. *et al. Nature* **309**, 724-726 (1984).
- Hirose, T., Crea, R. & Itakura, K. *Tetrahedron Lett.* **28**, 2449-2452 (1978).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161-170 (1982).
- Morrison, D. A. *Meth. Enzym.* **68**, 326-331 (1979).
- Wallace, R. B. *et al. Nucleic Acids Res.* **6**, 3543-3557 (1979).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Ohkubo, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2196-2200 (1983).
- McMaster, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835-4838 (1977).
- Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).

Heavily methylated amplified DNA in transformants of *Neurospora crassa*

John H. Bull & John C. Wootton

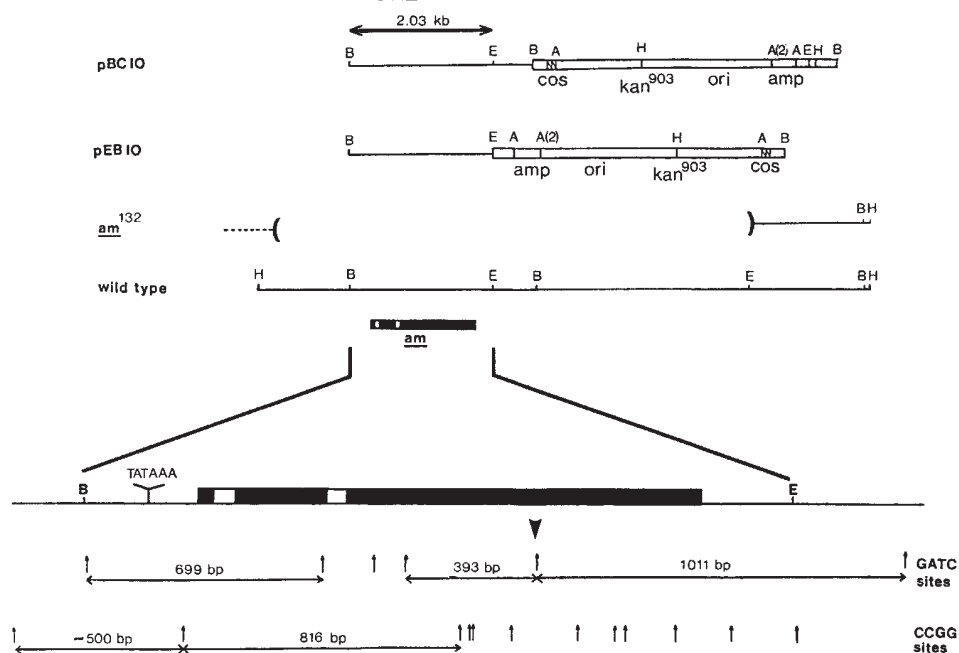
Department of Genetics, University of Leeds, Leeds LS2 9JT, UK

Substantial DNA methylation occurs in higher eukaryotes and in some cases affects gene expression (for reviews see refs 1-4). However, the genomes of some fungi^{5,6}, *Drosophila*⁷ and other lower eukaryotes^{8,9} have an extremely low 5-methylcytosine content, suggesting^{4,9} that DNA methylation might not have a general role in gene control. We have now found heavy methylation of transforming DNA which has become stably amplified in complex tandem arrays in the fungus *Neurospora crassa*. Rearranged amplified arrays of this type have not previously been found in fungi, but resemble those of animal systems¹⁰⁻¹⁶. Our results demonstrate that this lower eukaryote normally maintains only very low levels of 5-methylcytosine in its genome, but possesses a mechanism for substantial methylation of DNA *de novo*. This heavy methylation, which lacks the preference for CG sequences found in higher eukaryotes¹⁻⁴, does not apparently affect gene expression and might be involved in a recombination or repair process for which the amplified DNA is a target.

The transformation system used was unusual because the donor DNA (plasmids pBC10 and pEB10, Fig. 1) had no sequences in common with the recipient genome, but contained a selectable *N. crassa* gene, the cloned *am* gene encoding NADP-specific glutamate dehydrogenase¹⁷, which is commonly used in transformation experiments for *N. crassa* (refs 18, 19 and J.H.B. and J.C.W., unpublished). The *am*¹³² recipient has a long deletion that completely removes the *am* coding and flanking sequences (Fig. 1). Transformants of stable phenotype (*am*⁺ and resistant to high levels of antibiotic G418) were obtained by a two-stage selection procedure (see Fig. 2 legend for methods and frequencies). Southern blots (Figs 2, 3) probed with the donor DNA clearly showed multiple copies of the donor DNA in all nine transformants tested, thus contrasting with the single-copy transformants obtained with other *N. crassa* vectors (refs 18-22 and J.H.B. and J.C.W., unpublished). Copy numbers were stable and ranged from 4 to ~25, with the majority of transformants in the range 10-20 (Fig. 2).

The multiple copies of the donor DNA in these transformants are present in tandem arrays consisting of donor DNA that has undergone a complex set of rearrangements and partial deletions, shown by the many hybridizing fragments of different sizes given by all transformants tested in *Hind*III and *Aha*III digests (Figs 2, 3). The large size of hybridizing smears (Fig. 3) shows that segments of minimum length ~30-50 kilobases (kb) consist completely of DNA derived from the donor plasmids (original length ~7 kb, Fig. 1) and are not interspersed with recipient DNA or carrier DNA containing *Sal*I, *Kpn*I or *Hpa*I

Fig. 1 Restriction maps of the *am* region of the *N. crassa* genome and of the cosmids used for transformations. Cosmid vectors pBC10 and pEB10, shown in linear form, are aligned so that their segments of *N. crassa* DNA (single lines; bacterial vector sequences are shown as double lines) are above the corresponding part of the map of the wild-type *N. crassa* genomic DNA. *kan*⁹⁰³, the aminoglycoside phosphotransferase-3'(1) gene of bacterial transposon Tn903, which is expressed, conferring resistance to antibiotic G418, in yeast³² but which has not previously been reported to be expressed in *N. crassa*. The construction and properties of these vectors will be described elsewhere (J.H.B., N.H. Grimsley, J.C.W. and B. Hohn, in preparation). Black bars show the extent of the *am* gene (the two short introns are in white; transcription is left to right). The 9-kb region shown, of which the 2.7-kb *Bam*HI fragment containing the *am* gene has been sequenced¹⁷, is the *Hind*III insert originally cloned in a phage λ vector¹⁸. The *am*¹³² recipient has a deletion of at least 7 kb of linkage group V (shown above the wild-type map; the dashed line indicates that the left-hand limit is uncertain). pEB10 and pBC10 probes gave no detectable hybridization to blots of digested *am*¹³² DNA (Fig. 2, lane 10, Fig. 3, lane 9). Restriction sites: B, *Bam*HI; E, *Eco*RI; A, *Aha*III, A(2), two very close *Aha*III sites; H, *Hind*III. The enlarged region below the wild-type sequence shows the *Sau*3A–*Mbo*I (GATC) sites and *Msp*I–*Hpa*II (CCGG) sites predicted from the DNA sequence (upward-pointing arrows), and the sizes of the fragments detectable on Southern blots following cleavage with these enzymes. ∇ , The GATC site partially resistant to *Sau*3A cleavage (Fig. 4, lanes 13, 14) in wild-type DNA.



sites. Detailed mapping of these arrays will be reported elsewhere, as will meiotic analysis of the integration sites and the low fertility and 'barren' phenotype²³ of these transformants in crosses (J.H.B., J.C.W., Z. Nicola and B. Hohn, in preparation). Glutamate dehydrogenase assays showed that the level of expression of the *N. crassa am* gene in these transformants is surprisingly low, 5–20% of the normal enzyme levels of the wild-type strain, in spite of the multiple copies. Similar low levels of expression have previously been found in single-copy *am* gene transformants in cases in which integration was not at the normal *am* locus on linkage group V (refs 18, 19 and J.H.B. and J.C.W., unpublished).

The amplified transforming DNA is heavily methylated in most copies, as shown by Southern blots of transformant DNAs digested with the isoschizomer pairs *Sau*3A–*Mbo*I and *Hpa*II–*Msp*I (Fig. 4). Although a minor fraction of the transforming DNA was unmethylated, giving the small fragments (<1,500 base pairs (bp), Fig. 4) expected from the known sequences of the donor DNA, the major fraction of the amplified DNA was cleaved to large fragments (7–20 kb, Fig. 4) by the methylation-sensitive enzymes. Therefore, long segments of the amplified DNA, which include both bacterial and *N. crassa am* sequences, contained only 5-methylated cytosine in the many GATC and CCGG sites tested (also in CTGCAG sites, from *Pst*I digests not shown). The minor unmethylated fraction might represent all or part of some arrays, possibly occurring in different nuclei from arrays that are almost completely methylated, thus resembling 'compartments' of methylation found in some invertebrates²⁴ and chicken²⁵.

DNA of the wild-type strain of *N. crassa* contains only a very small amount of 5-methylcytosine, and ethidium bromide-stained gels (Fig. 4, lanes 9–12) showed indistinguishable distributions of fragment sizes between isoschizomers (except for some possibly significant uncleaved ribosomal RNA gene bands in the *Sau*3A digest; Fig. 4, lane 9). This was the case for the wild-type strain, the *am*¹³² recipient and all the transformants tested. Therefore, only the amplified DNA in these transformants is heavily methylated; in the rest of the genome, at most

only a very low level of 5-methylcytosine is maintained, as in the normal wild-type genome. A 1,404-bp fragment (Fig. 4, lane 13) is present in addition to the products of complete digestion, providing evidence for a minor amount of methylation at one *Sau*3A–*Mbo*I site in the normal *am* gene of the wild type (indicated in Fig. 1). This is a GATCG site, which therefore meets the symmetry requirements (for CG or CNG) arising from the action of maintenance methylase mechanism on hemimethylated DNA, as proposed for higher eukaryotes^{1–4}. In contrast, there is little sequence specificity in the heavy methylation of the amplified arrays, shown by the unbiased occurrence of all four nucleotides adjacent to the GATC sites in the *N. crassa am* DNA probed (sites deduced from the known sequence are shown in Fig. 1) and in flanking bacterial DNA of known sequence. Also, this DNA was almost equally resistant to cleavage by *Msp*I and *Hpa*II (Fig. 4), indicating a great majority of doubly methylated (mCmCGG) rather than singly methylated sites.

Clearly, the initial methylation in *N. crassa* of the donor transforming DNA, which was purified from *Escherichia coli*, was *de novo*, and the proportion of methylated arrays continued to increase during vegetative growth (Fig. 4, lanes 17–22 show results from mycelial samples at different times). Growth in the presence of 300 μ M 5-azacytidine completely prevented this methylation (Fig. 4, lanes 15, 16), but had no significant effect on the level of expression of the *N. crassa am* genes present in the amplified DNA, as measured by glutamate dehydrogenase assays. This suggests that the heavy *de novo* methylation is not involved in the control of gene expression, but rather must reflect the presence in *N. crassa* of a mechanism distinct from a conventional maintenance methylase. We suggest that the irregular sequence repeats of the amplified DNA are likely to be hotspots for the action of recombination processes and repair synthesis, one step of which might involve methylation of cytosine *de novo*. 'Repair methylation' has been observed in mammalian cells following DNA damage²⁶ or S-phase arrest²⁷, and involves methylation specifically of cytosine inserted into DNA strands by repair synthesis. An association of methylation

Fig. 2 Multi-copy DNA in transformants of *N. crassa*. Digested DNA samples were from the following strains (lane numbers in parentheses): wild-type *N. crassa* STA4 (1); untransformed *am*¹³² recipient (10); transformants E222 (3); BC3 (2); E223 (5, 8); E224 (4, 7); ECT845 (6); EB847 (9); EB841 (11). The code letters of the transformants denote the donor DNA: E, EB or ECT for pEB10, BC for pBC10. Restriction enzymes used were (lane numbers in parentheses): *Aha*III (1–5); *Hind*III (6–11). Vertical bars show the positions and sizes of phage λ DNA *Hind*III fragments run as markers on the same gels.

Methods: Transformation of sphaeroplasts of recipient strain *am*¹³² *ini*³⁷⁴⁰¹ (ref. 33) was carried out essentially as described in ref. 34, using in each experiment $\sim 10^8$ conidia in 0.4 ml transformation buffer mixed with 20 μ g of donor cosmid DNA and, in some experiments, 20 μ g of carrier DNA (either calf thymus DNA or total cellular DNA prepared from the *am*¹³² strain; carrier DNA did not significantly affect the transformation frequency and was omitted in later experiments). After 30 min incubation on ice and 30–45 min incubation at 22 °C (polyethylene glycol additions were as in ref. 34), 10-fold dilutions of treated sphaeroplasts were plated in an overlay of Vogel's minimal medium³⁵ containing 1 M sorbitol, 1.5% w/v L-sorbose, 0.05% w/v of each of D-glucose and D-fructose and 3% w/v agar³⁴, supplemented with 20 mM glycine (to suppress the leaky growth of untransformed *am*¹³² clones) and 0.1% w/v myo-inositol. After incubation for 5 days at 30 °C, frequencies of *am*⁺ transformants were 3.5 per μ g of cosmid DNA, 7.8×10^{-6} per viable sphaeroplast for pBC10 DNA and 1.8 per μ g of cosmid DNA, 5.2×10^{-6} per viable sphaeroplast for pEB10 DNA (averaged results from five separate experiments; abortive and dubious slow-growing transformants were excluded from these calculations). These frequencies are ~ 10 –20-fold less than those obtained in similar experiments using other cosmids containing the 9-kb *Hind*III fragment encompassing the *am* gene (Fig. 1) which yield single-copy transformants (J.H.B. and J.C.W., unpublished). *am*⁺ clones were transferred to plates of Vogel's/glycine/inositol medium (as above but lacking sorbitol and containing 1.5% w/v agar) containing 2 mg ml⁻¹ antibiotic G418 (obtained from P. J. L. Daniels, Schering, USA). After 7–12 days incubation at 30 °C, G418-resistant clones grew from 85% of the *am*⁺ inocula, and these were maintained as stocks on both selective and non-selective media. Stability of the G418^r phenotypes was shown by resistance tests of duplicate stock cultures maintained with or without G418 on medium selective for *am*⁺. The transformants showed various characteristic maximum tolerance levels to G418, ranging over 2–10 mg ml⁻¹ compared with the parental *am*¹³² strain which is sensitive to 0.5 mg ml⁻¹. DNA for Southern blot analysis was purified from mycelial cultures (grown in 50 ml of liquid Vogel's minimal medium supplemented with 20 mM glycine for 2 days) of an arbitrarily chosen set of nine of the *am*⁺ G418^r transformants (representing five separate transformation experiments), by the rapid methods of refs 36, 37, except that agitation was avoided during extraction and all volumes were scaled down proportionally to fit operations into 1.5-ml Eppendorf tubes. 0.8% agarose gels were loaded with equal amounts (4 μ g per slot) of DNA samples digested to completion with 10 units of restriction enzymes per 4 μ g DNA for 4–6 h. Completeness of digestion was judged from the characteristic size distribution of fragments on ethidium bromide-stained gels and the appearance of only the well characterized, completely digested bands of ribosomal RNA genes. Following transfer to nitrocellulose filters, hybridization and washes were in standard high-stringency conditions³⁸ using pEB10 DNA as probe, ³²P-labelled by nick-translation to at least 2×10^8 c.p.m. per μ g. Copy numbers were estimated by comparison with the single-copy bands of wild-type *N. crassa* DNA (lane 1 and Fig. 3, lane 16) using microdensitometric scans of the autoradiographs shown, of different exposure times of the same filters, and also of other similar filters (not shown) probed separately with labelled restriction fragments containing the *am* or *amp* regions of the donor plasmids. Stability of copy number was shown by blot hybridization of separate DNA preparations from mycelium at different stages of vegetative growth and from cultures inoculated from separate stocks maintained on selective and non-selective media. In all such cases the pattern of hybridizing bands was essentially identical to those shown.

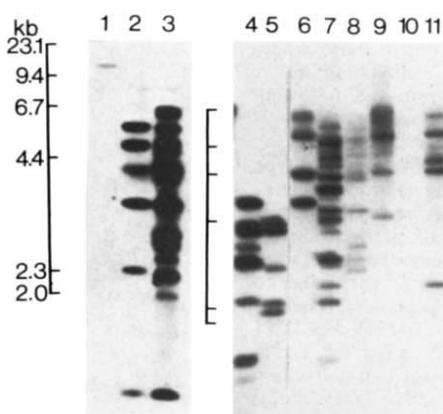


Fig. 3 Amplified transforming DNA is present in tandem arrays of high molecular weight. DNA samples were from the following strains (lane numbers in parentheses): wild-type STA4 (16); untransformed *am*¹³² recipient (9); transformants (coded as in Fig. 2) EB3 (1, 2); E221 (3, 4, 13); E222 (5, 6, 14, 15); ECT845 (7); EB847 (8); EB841 (10); E224 (11); E223 (12). Restriction enzymes used were (lane numbers in parentheses): *Sal*I (1, 3, 5, 7–12); *Hind*III (2, 4, 6); *Aha*III (13); *Kpn*I (14); *Hpa*I (15); *Bam*HI (16). Probes were pBC10 (lanes 1–6) and pEB10 (lanes 7–16). Markers were as in Fig. 2.

Methods: Methods for DNA preparation and Southern blot analysis were as described in Fig. 2 legend.

Completeness of restriction digestion was shown by ethidium bromide-stained gels which gave the distribution of genomic fragments typical of *N. crassa* in all cases. The minimum size of the tandem arrays of amplified DNA was estimated from the Southern blots of transformant DNA digested with restriction enzymes (*Sal*I, *Kpn*I and *Hpa*I) whose recognition sites are absent in the donor plasmids pBC10 and pEB10. In these digests the hybridizing DNA gives a high-molecular weight smear in size range (20–50 kb) typical of *N. crassa* DNA prepared by the rapid methods^{36,37} used. A single exception was transformant ECT845 (lane 7), which shows two *Sal*I fragments (22 and 20 kb) and might either have incorporated some calf thymus carrier DNA into an array or have two relatively small arrays flanked by recipient DNA. Rescue of the amplified arrays into *E. coli* by packaging *in vitro* in phage λ heads (to be reported elsewhere, J.H.B., J.C.W., Z. Nicola and B. Hohn, in preparation) gave further evidence for the tandem arrangement of the multiple copies of donor DNA, as the packaging reaction requires pairs of *cos* sites spaced at 38–52 kb³⁹.

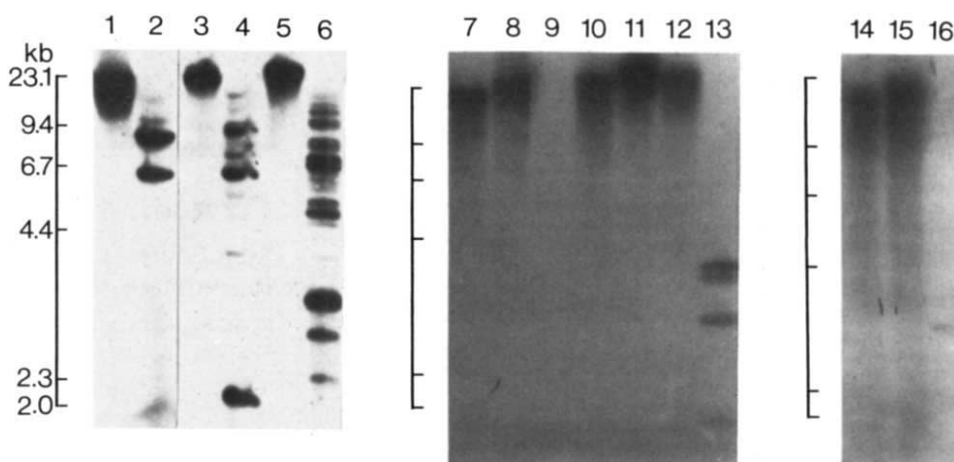
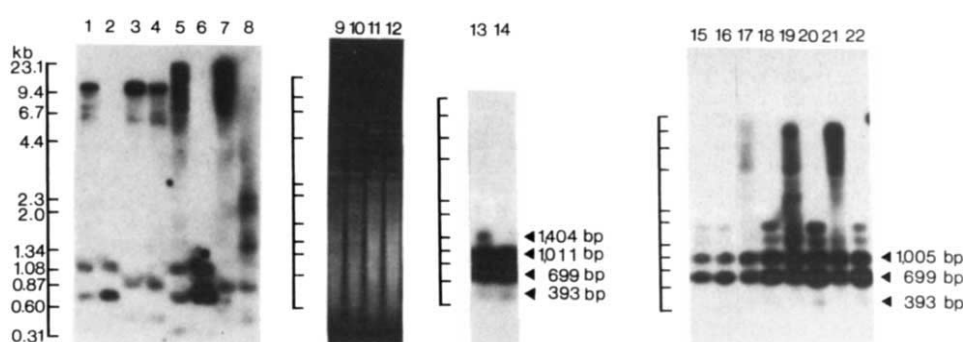


Fig. 4 Amplified transforming DNA is heavily methylated and 5-azacytidine prevents methylation. DNA samples were from the following strains (non-standard growth conditions are indicated if relevant; lane numbers in parentheses): wild-type STA4 (13, 14); the hybridizing bands of wild-type DNA were ~15-fold less intense than those of the transformant DNAs, but the autoradiographs were exposed for longer to give a similar appearance; transformants BC3 (1–4); E222 (5–12); E221 (15, 16); DNA prepared from mycelium grown for 16 h in medium containing 300 μ M 5-azacytidine; E222 mycelium from 12 h liquid culture (17, 18); E222 mycelium from 24 h liquid culture (21, 22). Restriction enzymes used were (lane numbers in parentheses): *Sau*3A (1, 5, 9, 13, 15, 17, 19, 21); *Mbo*I (2, 6, 10, 14, 16, 18, 20, 22); *Msp*I (3, 7, 11); *Hpa*II (4, 8, 12). Lanes 9–12 show the ethidium bromide-stained gel from which the blot of lanes 5–8 was made. For all Southern blots the probe was the 2.03-kb *Bam*HI–*Eco*RI fragment of *N. crassa* DNA encompassing the *am* gene (Fig. 1). Size markers (vertical bars) were a mixture of phage λ *Hind*III fragments and bacteriophage Φ X174 *Hae*III fragments. Arrows indicate the sizes of hybridizing fragments expected from the sequenced region probed (see the map of GATC sites indicated in Fig. 1) and the expected 1,005-bp boundary fragment that includes bacterial sequences. The 1,404-bp fragment marked (lane 13) results from the failure of *Sau*3A to cleave the GATC site between the 393- and 1,011-bp fragments. Other clearly defined fragments of ~1,300 and 1,600 bp (from E221 and E222 DNA, lanes 15–22) are probably new boundary fragments generated by deletion or rearrangement of DNA during amplification.



Methods: Standard medium for liquid cultures was Vogel's minimal medium (Fig. 2 legend) containing 20 mM glycine and 0.1% w/v *myo*-inositol. 300 μ M 5-azacytidine was added as indicated; preliminary experiments showed that 3 μ M 5-azacytidine, as commonly used to inhibit DNA methylation in animal cells¹⁻⁴, had no effect in *N. crassa*. Methods of DNA preparation and Southern blot analysis, and criteria for completeness of restriction digestion were as described in Fig. 2 legend. DNA methylation was tested by means of the isoschizomer pairs *Sau*3A (which does not cleave GATmC) and *Mbo*I (which cleaves GATmC)⁴⁰ and *Hpa*II (which does not cleave CmCGG or mCmCGG but cleaves mCCGG) and *Msp*I (which cleaves CmCGG, except in GGmCGG sequences⁴¹, but does not cleave mCCGG or mCmCGG, ref. 40) where mC denotes 5-methylcytosine. Completeness of *Mbo*I digests shows that N⁶-methyladenine is not present at GATC sites.

with repair has also been suggested for UV-irradiated bacteria²⁸ and some recombinogenic lesions of 'hyper-rec' strains of *E. coli*²⁹. In animal systems, DNA amplification, a frequent mechanism by which cells meet demands for increased amounts of gene products¹⁰⁻¹⁶, does not generally¹⁰ involve abnormally heavy methylation. However, *de novo* methylation of transfected foreign DNA sequences³⁰ has been found, and heavy hypermethylation of amplified rRNA genes occurs in rat hepatoma and some human cell lines³¹ which, together with our results, suggests that methylation *de novo* of some types of amplified or rearranged DNA might be a widely distributed phenomenon.

We thank the SERC (research grant GR.B.34423 to J.C.W.) for financial support.

22. Buxton, F. P. & Radford, A. *Molec. gen. Genet.* **190**, 403–405 (1983).
23. Perkins, D. D. & Barry, E. C. *A. Rev. Genet.* **11**, 133–225 (1977).
24. Bird, A. P. & Taggart, M. H. *Nucleic Acids Res.* **8**, 1485–1497 (1980).
25. Cooper, D. N., Taggart, M. H. & Bird, A. P. *Nucleic Acids Res.* **11**, 647–658 (1983).
26. Drabovsky, D., Lacko, I. & Wacker, A. *Biochim. biophys. Acta* **447**, 139–143 (1976).
27. Hilliard, J. K. & Snider, T. W. *Nucleic Acids Res.* **2**, 809–819 (1975).
28. Ryan, A. M. & Borek, E. *Biochim. biophys. Acta* **240**, 203–214 (1971).
29. Korba, B. E. & Hays, J. B. *J. molec. Biol.* **157**, 213–235 (1982).
30. Pollack, Y., Stein, R., Razin, A. & Cedar, H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6463–6467 (1980).
31. Tantravahi, U., Guntaka, R. V., Erlanger, B. F. & Miller, O. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 489–493 (1981).
32. Jimenez, A. & Davies, J. *Nature* **287**, 869–871 (1980).
33. Wootton, J. C., Fraser, M. J. & Baron, A. J. *Neurospora Newslett.* **27**, 33 (1980).
34. Schweizer, M., Case, M. E., Dykstra, C. C., Giles, N. H. & Kushner, S. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5086–5090 (1981).
35. Vogel, H. J. *Am. Nat.* **98**, 435–440 (1964).
36. Metzner, R. L. & Baish, T. J. *Neurospora Newslett.* **28**, 20–21 (1981).
37. Stevens, J. N. & Metzner, R. L. *Neurospora Newslett.* **29**, 27–28 (1982).
38. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
39. Hohn, B. *Meth. Enzym.* **68**, 299–309 (1979).
40. McClelland, M. *Nucleic Acids Res.* **11**, r169–r173 (1983).
41. Busslinger, M., de Boer, E., Wright, S., Grosveld, F. G. & Flavell, R. A. *Nucleic Acids Res.* **11**, 3569–3571 (1983).

Received 5 April; accepted 21 June 1984.

1. Razin, A. & Riggs, A. D. *Science* **210**, 604–610 (1980).
2. Ehrlich, M. & Wang, R. Y.-H. *Science* **212**, 1350–1357 (1981).
3. Adams, R. L. P. & Burdon, R. H. *CRC Crit. Rev. Biochem.* **13**, 349–384 (1982).
4. Doerfler, W. A. *Rev. Biochem.* **52**, 93–124 (1983).
5. Hattman, S., Kenny, C., Berger, L. & Pratt, K. J. *Bact.* **135**, 1156–1157 (1978).
6. Feher, Z., Kiss, A. & Venetianer, P. *Nature* **302**, 266–268 (1983).
7. Achwal, C. W., Ganguly, P. & Chandra, H. S. *EMBO J.* **3**, 263–266 (1984).
8. Urieli-Shoval, S., Gruenbaum, Y., Sedat, J. & Razin, A. *FEBS Lett.* **146**, 148–152 (1982).
9. Bird, A. P. *Nature* **307**, 503–504 (1984).
10. Schimke, R. T. *Gene Amplification* (Cold Spring Harbor Laboratory, New York, 1982).
11. Beach, L. R. & Palminter, R. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2110–2114 (1981).
12. Spradling, A. C. *Cell* **27**, 193–201 (1981).
13. Dalla-Favera, R., Wong-Staal, F. & Gallo, R. C. *Nature* **299**, 61–63 (1982).
14. Schwab, M., Alitalo, K., Varmus, H. E. & Bishop, J. M. *Nature* **303**, 497–501 (1983).
15. Roberts, J. M. & Axel, R. *Cell* **29**, 109–119 (1982).
16. Roberts, J. M., Buck, L. B. & Axel, R. *Cell* **33**, 53–63 (1983).
17. Kinnaird, J. H. & Fincham, J. R. S. *Gene* **26**, 253–260 (1983).
18. Kinnaird, J. H., Keighren, M. A., Kinsey, J. A., Eaton, M. & Fincham, J. R. S. *Gene* **20**, 387–396 (1982).
19. Kinsey, J. A. & Rambousek, J. A. *Molec. cell. Biol.* **4**, 117–122 (1984).
20. Case, M. E., Schweizer, E. M., Kushner, S. R. & Giles, N. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5259–5263 (1979).
21. Case, M. E. in *Genetic Engineering of Microorganisms for Chemicals* (ed. Hollaender, A.) 87–100 (Plenum, New York, 1981).

Erratum

Functions of the canal system in the rotaliid foraminifer, *Heterostegina depressa*

R. Röttger, M. Spindler, R. Schmaljohann, M. Richwien & M. Fladung

Nature **309**, 789–791 (1984)

ON page 789, the last sentence of the second paragraph should read: We have examined here the symbiont-bearing Recent nummulitid, *H. depressa*, in which the canal system located within the test walls replaces the primary aperture seen in other foraminifera.

The computer experience

David Edge

Turing's Man: Western Culture in the Computer Age.

By J. David Bolter.

University of North Carolina Press: 1984. Pp.264. Hbk \$19.95, pbk \$8.95.

DAVID Bolter teaches classics at North Carolina, and has been a visiting fellow in computer science at Yale. In *Turing's Man* he offers us a bold interdisciplinary study of "the impact that electronic logic machines are having upon our culture", concentrating on the "subtler effect" of "a change in the way men and women in the electronic age think about themselves and the world around them" (p.4). Computers are a "defining technology" (p.8): they direct and reorganize older technologies, which are then rendered subservient; and, in so doing, they reshape the central experiences of their age. They therefore lend themselves to metaphorical extension, leading us to reconceptualize the world in their own image — much as some earlier technologies (potters' wheels, clocks, steam engines) did before them. "By promising (or threatening) to replace man, the computer is giving us a new definition of man as an 'information processor', and of nature as 'information to be processed'" (p.13). Bolter calls those who hold this view, Turing's Men.

Following a brief historical survey of the successive cultural impacts of manual, mechanical and dynamic technologies, the bulk of the book is taken up with a popular account of the principles of electronic computers (which I am not convinced is all either necessary or entirely successful for the uninitiated, but is still impressive), followed by chapters devoted to the implications of computer technology for mathematics and logic, and for our conceptions of space, time and progress, language, memory, creation and intelligence. Turing's Man "thinks of his world, intellectual and physical, as finite" (p.226), emphasizing discrete arrays rather than continua; lives, like the Greeks, for the moment, lacking historical sensibility, and seems "destined to lose the Faustian concern with depth" (p.220); and, since his interaction with computers is essentially playful, realizes that games are deadly serious. Even "the Western concept of God as an infinite being must surely fade" (p.226). You can't really be bolder than that!

Perhaps realizing that his argument must be self-exemplifying, in his preface Bolter apologizes for his necessarily superficial and general approach. But the discussion of the subject must begin somewhere, and Bolter's first steps are as good as any; his book is a readable and stimulating introduction to a profound intellectual issue. If the debate is to proceed, however,

some clarification of the argument is needed.

To begin with, is this another version of technological determinism? Is Turing's Man the necessary result of the development of inherent properties of a physical technology — a development which is inevitable, inexorable, irreversible and beyond social control? Like many other

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Turing's children — "destined to lose the Faustian concern with depth"?

authors, Bolter seems ambivalent on this point. His early discussion of the emergence of computers stresses the importance of economic conditions in stimulating their development, but he argues that these conditions "enabled these devices to express qualities that were latent in them from the moment the first prototypes were tinkered with" (p.5) — and it is from these latent qualities that the characteristics of Turing's Man stem. Can such "hard" aspects be avoided? In his closing discussion, Bolter argues that the tendencies he has described "need not overwhelm us" (p.228); but his strategy for "building humaneness into the machine" (p.229) centres entirely on the programming, the software; he seems to take the hardware as an unavoidable "given". But it is precisely such an implicit technological determinism that many of us find questionable. The hardware itself may be socially shaped. The notion that culture adapts to, and stems from, the inevitable unfolding of the latent characteristics of physical technology is a notion that demands further challenge and analysis.

Secondly, Bolter assumes a cultural

homogeneity. The Greek artisans and craftsmen are silent: but Plato and Aristotle speak for them. The potters, weavers and carpenters struggle with their "defining technologies", experiencing and conceptualizing the world in their terms: and the artists, playwrights and speculative thinkers reproduce that world for our inspection. I find this assumption hard to accept. There must surely be a very wide variation, within any society, of experience of the "defining technologies". Turing's Man is essentially a designer and programmer of computer technology — someone who manipulates and creatively explores its potential. Will this kind of experience ever be available to more than a minority? Just as we are familiar with cars and TV sets without designing, mending, tinkering with, or even understanding them, so we may become familiar with computers without sampling the experiences which beget Turing's Men. Even the elementary popular understanding which Bolter presents may be redundant. Why, then, should Turing's Man be taken as representative of an entire culture? Must we all conform to this pattern?

This question leads us back to my point about technological determinism. Just as it is possible that the precise physical form of the technologies we develop are shaped by powerful social interests, so it may be that the ways in which such technologies are metaphorically (conceptually, culturally) extended are also socially moulded. In other words, the reconceptualizations may serve ideological functions, redefining the world so that developments sponsored by a minority come to seem natural and inevitable. Rewriting history with just the same kind of broad brush that Bolter uses, the emergence of mechanical clocks, and their extension into a Newtonian metaphysics, is a story that can plausibly be told in such terms. However, my point is not to argue over particular historical episodes. Rather, it is to stress the importance of increasing our understanding of the social processes that underly *both* the evolution of physical technologies *and* their cultural extension.

Bolter attempts no such analysis. His book is an extended meditation on a particular technological metaphor. To regret that it does not do justice to the relevant literature on metaphor (and, in particular, to the pioneering work of Donald Schon) is, perhaps, a Faustian sentiment which we must now eschew. And many of Bolter's speculations seem strained and implausible. But no matter. His concern, as he says, "is not that the reader agree with all my conclusions but rather that he or she agree that it is important to think about computers from this perspective" (p.xii). And indeed it is. □

David Edge is Reader in Science Studies and Director of the Science Studies Unit at the University of Edinburgh.

Camera Press

Ecological models and mechanisms

Richard G. Wiegert

Systems Ecology: An Introduction.

By Howard T. Odum.

Wiley: 1983. Pp.644. \$54.95, £56.25.

IN *Systems Ecology*, Howard T. Odum attempts to wrap virtually all living systems in the explanatory mantle of systems science and energetics. The mantle is frayed in some places, stretched too thin in others and is certainly far from large enough to cover everything, yet credit must go to the author for simply making the attempt. Odum's arguments touch, in one manner or another, on just about every aspect of ecology. They are often controversial, occasionally wrong, rarely naive but never dull.

The book consists of 27 chapters, divided into four sections, the first of which is a general overview of systems ecology and modelling. The remaining three deal with the main elements and interactions found in ecological models, the organization of systems in terms of energy and other physical factors, and the system attributes of various levels of biological organization. This last section concludes with a chapter on the unity of systems.

In the first part Odum makes a laudable attempt to describe and compare the various possible ways (that is, the languages) that can be used to explain the workings of a system and translate them into forms understandable by a computer. Unfortunately, the set of defined "languages" he discusses (given in Table 1-1) is, in the sense of comparing apples and oranges, a veritable fruit-basket. The language of mathematical notation, for example, has been subdivided into a number of "different" languages before being compared to English or the energy-circuit language long-used effectively by Odum. Many workers choose to use the latter in its entirety. Others, myself included, use mathematical notation, perhaps supplemented by some of the more general state-variable symbols of Forrester or Odum. A much more careful and rigorous definition of "language" would have immensely increased the usefulness of the comparison.

In Part II Odum systematically lists the functional descriptions of ecological interactions that have been used in the past. These cover control of autocatalytic modules, influence loops, serial predator-prey interactions, parallel competition interactions and food webs. This is a most helpful entrée to the literature. But the effort to find and convert all such functions to the language of energy circuits has unfortunately left no room for assessment of the various functions and their relevance in light of theory. For example, the important

topic of r versus K strategy is disposed of in a single paragraph of eight lines!

Part III deals with the importance of energy in system organization and pattern. In Odum's concept of "embodied energy", the value of the energy content of matter is augmented in proportion to its position or distance (in terms of trophic transformations) from the Sun. This is an important distinction. But I believe it is overdone by drawing an analogy with temperature, as the latter is used to measure the "availability" of heat energy. Temperature is an exact reference because it always has an absolute (zero degrees Kelvin) or a relative reference (the temperature of the surroundings). But energy transformations in living systems involve many unknown pathways and are variable; organisms often feed from several trophic levels while accumulating their stored energy. Furthermore, the human value-system may often make the concept of embodied ecological energy meaningless. For example, the pearl has the same embodied ecological energy as an equal weight of oyster shell, yet, because of its value, it can affect ecosystems to a much larger degree.

In general this part of the book suffers from too great an emphasis on physical examples. The chapter on temperature would have profited from mention of the temperature functions in models necessitated by the great wealth of behavioural and physiological adaptations of organisms to ameliorate the effects of varying temperature. More importantly, Odum is wrong when he says (p.318) that "The energy flow of maintenance respiration

divided by the Kelvin temperature gives the entropy change". This old analogy presumes the living entity in question to be closed to exchanges of matter and in reversible thermodynamic equilibrium. Neither is true and the entropies of exchanging matter and the irreversible heat loss completely swamp the rather small entropy changes associated with internal chemical transformations. This is why we use First Law energy balance equations in ecological energetics.

In Part IV Odum discusses system modelling of successively larger ecological hierarchies, culminating in a presentation of world patterns. This, the freshest and most interesting part of the book, contains more detailed and free-ranging discussion and critical evaluation. In particular the modelling of monetary flows as well as of energy flows is explained clearly and there are some excellent examples of embodied energy expressed in monetary terms.

Alongside the stress on a holistic view of ecosystems, throughout the book Odum implicitly makes a strong case for modelling ecosystems with components for which mechanistic explanations have been supplied. This in turn implies an awareness of and reliance upon population genetic selection and fitness as the basic units of evolution in ecosystems. The incorporation of these concepts into ecosystem models is a prime objective of most ecological modellers today. Explicit treatment of this subject in the next edition of the book would be a valuable addition. □

Richard G. Wiegert is Professor of Zoology at the University of Georgia.

Star appeal

R. J. Tayler

Superstars: How Stellar Explosions Shape the Destiny of Our Universe.

By David H. Clark.

McGraw-Hill: 1984. Pp.216. \$17.95.

WHEN a star explodes as a supernova, there is a short period in which it is as luminous as a whole galaxy of ordinary stars. The sudden appearance of such a "new star" has in the past been regarded as a portent, and two supernovae observed in 1572 and 1604 played a part in the acceptance of the new astronomy in which all objects beyond the Moon were no longer regarded as perfect and unchanging. Supernovae are important sources of heavy chemical elements and of cosmic rays; their explosions are crucial to the energy balance of the interstellar medium; they may stimulate the formation of new stars; and, in some cases, their explosions leave behind highly compact and exotic remnants, pulsars. It has also been suggested that biological extinctions might have been caused by the occurrence of supernovae

relatively close to the Earth.

David Clark gives an account of all these aspects of supernovae for a general audience. He is at his best when discussing the historical significance of supernovae and describing how, for example, the Chinese and Korean astronomical records have provided evidence for the explosions of supernovae whose remnants can be studied today. Although the book also contains a good assessment of the astronomical importance of supernovae without the introduction of any mathematics, it inevitably contains many complicated physical ideas, some but not all of which are conveniently discussed in appendices. Because of this the book will be most readily appreciated by readers with a basic training in physics.

Anyone who does read the book will realize why astronomers are hoping that there will soon be another bright supernova in our Galaxy. Every possible instrument would be used to study the supernova and the early development of its remnant, and the increase in astronomical knowledge would potentially be very great indeed. □

R.J. Tayler is Director of the Astronomy Centre of the University of Sussex.

Feeding and ranging in the forest

R.D. Martin

Five New World Primates: A Study in Comparative Ecology.

By John Terborgh.

Princeton University Press: 1984. Pp. 260. Hbk \$40, £40; pbk \$13.50, £10.40.

ADDITIONS to the literature on New World monkey behavioural ecology are rare enough, so a monograph dealing with five species at once is particularly newsworthy. In 1976 John Terborgh, originally an ornithologist but now a convert to primatology, opened up a field site in pristine Amazonian rain forest on the Rio Manu in Peru, and in *Five New World Primates* he has summarized a fascinating array of behavioural and ecological observations. As the first contribution to a new series, *Monographs in Behavior and Ecology*, the book sets a standard that augurs well for the future. It is essential reading for anyone interested in primate behaviour and ecology, or in the subject of behavioural ecology in a wider sense.

Terborgh's account deals in detail with five out of the unusually high total of 11 primate species found at the Manu site, and includes useful information on the other six. Although the fieldwork spanned little more than a year, a demanding rota of observations shared with four assistants permitted a remarkable weekly average of 50 hours' systematic data collection for the five main species: two capuchins (*Cebus apella* and *C. albifrons*), two tamarins (*Saguinus imperator* and *S. fuscicollis*) and a squirrel monkey (*Saimiri sciureus*). The author modestly describes much of the work as "first-order descriptive natural history", yet even at that level the information provided is of great value. However, the text goes far beyond description

because of the emphasis on standardized collection of quantitative data on feeding and ranging behaviour. Further, there are special advantages in conducting a study of several species simultaneously — comparisons between sympatric species are particularly productive because many environmental variables are held constant, and it is also possible to obtain quite precise information on divergent specializations that reduce competition between species.

Terborgh's data on feeding and ranging behaviour permit construction of detailed profiles of the five species studied, the profiles in turn providing a basis for comparison both between these species and with other primates. Overall, the picture built up by Terborgh reveals a stark contrast with the better-known Old World monkey species. For example, the Manu primates feed far more extensively on animal prey (notably arthropods), whereas leaf-eating is insignificant. There are also quite substantial differences in feeding behaviour between the five species, these dietary differences being matched by general differences in habitat use. There is, therefore, good evidence for some avoidance of competition between species.

However there is also some overlap between the five species in their feeding activities and thus some observable competitive interaction (especially between the tamarin species). It is therefore all the more striking that there are two special cases of mixed troop formation (polyspecific association), which further enhance the importance of Terborgh's work. First, there is a definite tendency for *Saimiri* to form mixed troops with either of the *Cebus* species. This has been reported previously, but Terborgh's quantitative data show quite clearly that it is *Saimiri* rather than *Cebus* that seek the association. For instance, whereas *Saimiri* travel relatively rapidly when alone and slow down when in association with *Cebus* (often after switching direction to make contact), the converse is true of the capuchins. Terborgh rejects a suggestion that *Cebus* benefits from the association by preying on insects flushed by *Saimiri* and concludes instead that *Saimiri* may benefit through greater protection from predators and through the more detailed local knowledge of fruit sources presumably possessed by the less widely ranging *Cebus*.

The second case of troop association is even more noteworthy in that an *S. imperator* group typically shares a territory with a group of *S. fuscicollis*, while both groups vigorously repel conspecific intruders. Unlike *Saimiri* and *Cebus*, however, the two tamarin species do not actually move as mixed troops, though their ranging patterns are closely coordinated. In this case, Terborgh invokes Cody's "local depletion hypothesis" (originally proposed for mixed flocks of birds), according to which the two tamarin species increase foraging efficiency by avoiding previously depleted areas. It is, however, odd that the tamarins

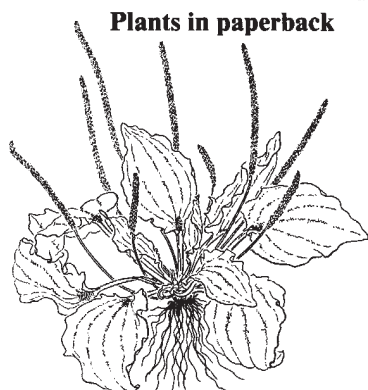
diverge in their patterns of feeding on arthropod prey, whereas they are extremely similar in their feeding on fruit. Further, Cody's hypothesis does not seem to resolve the basic paradox that arises with any mixed troop: why does a group of species A associate with a group of species B rather than forming a larger single-species group?

Inevitably, such an ambitious study has its shortcomings. Because of the overriding need to standardize observation methods, focal animal sampling (requiring recognition of individuals) was ruled out and replaced by the less rigorous technique of opportunistic scan sampling. This latter method predisposes towards recording the most obvious activities of the most obvious individuals. The reasons for adopting scan sampling are understandable, but the potential effects of observational bias on inferred activity budgets really required some detailed discussion. Also, although Terborgh and his team took pains to assess fruit distribution and productivity at the Manu site, there was no comparable assessment of arthropod availability. This is undoubtedly linked to Terborgh's general tendency to dismiss competition for arthropods as a determining influence in relationships between the five species, in favour of arguments based on fruit availability and on risks of predation. For this and other reasons, Terborgh's overall attempts to interpret the behaviour of the Manu primates in relation to ecological factors are not entirely convincing.

Finally, although Terborgh draws upon the ornithological literature in interpreting the behavioural ecology of the Manu primates, there is surprisingly little reference to directly comparable research on Old World sympatric primate species — for example that of the Gautiers on monkeys in Gabon (including the best examples known of polyspecific association among primates) and of Chivers *et al.* on monkeys and lesser apes in Malaysia, of Charles-Dominique on nocturnal prosimians in Gabon, and of Charles-Dominique *et al.* on nocturnal prosimians in West Madagascar. All of these studies establish general principles of direct relevance to Terborgh's work at Manu, but it is left to the reader to make the comparisons and the connections.

However such points pale into insignificance given the dedication required to carry out such a field study at all. The introductory account of the practical problems involved is enough to arouse the reader's unstinted admiration. In this book Terborgh has greatly extended our understanding of the behavioural ecology of Amazonian primates. As noted in the epilogue, one of the benefits of such understanding may be the development of effective conservation measures for primates in this fast-disappearing rain forest habitat. □

R.D. Martin is Professor of Physical Anthropology at University College London.



Great Plantain (*Plantago major*), depicted by Sybil J. Roles. The drawing is reproduced from the new paperback edition of *Flora of the British Isles, Illustrations*, the companion volumes to Clapham, Tutin and Warburg's classic work. The illustrations are published in four volumes by Cambridge University Press, price £8.95, \$17.95 each volume.

Not yet the meeting of minds

John Morton

Handbook of Cognitive Neuroscience.

Edited by Michael S. Gazzaniga.

Plenum: 1984. Pp.416. \$45, £34.65.

COGNITIVE science is a ten-year-old conglomerate which is still struggling for identity, survival and research funds. Its scope, as the term "cognitive" indicates, includes perception, memory, language, intelligence and reasoning. In general it excludes such topics as motivation, intention, feeling and emotion. The core disciplines are cognitive psychology, theoretical linguistics and artificial intelligence. The links between the three are still rooted in optimism rather than achievement, and the merger with the neurosciences, to form cognitive neuroscience, is a triumph of faith.

The urge towards a unified science of mind is understandable, but the success of a cooperative enterprise depends on the existence of a question that is susceptible to a common interpretation. This is rarely the case. Thus, workers in artificial intelligence often study the same subject matter as cognitive psychologists — for example the problem of object recognition — but the approach is one of finding *any* solution rather than the specifically human solution. When we include the neurosciences there are difficulties of another sort. Given that we believe a particular psychological mechanism exists then we can ask how it is implemented in the brain. But, unless we are attempting causal accounts, it is not clear that there are any constraints on psychological theory from biology except where there is one-to-one mapping between the elements at the two levels. From what is currently believed, such a mapping exists at best for sensory and motor functions.

In practice, then, each cognitive scientist (self-defined) gets what he can from the other disciplines in pursuit of his own objectives. This is profitable in that it can lead to an enrichment of the metaphorical and analogical resources available to the individual. But it is not the same as the creation of a unified science.

By now the reader should have some idea of what to expect from a *Handbook of Cognitive Neuroscience*. Michael Gazzaniga, the editor, is a psychologist best known for his insightful studies of split-brain patients. He has gathered together a collection of 19 articles which well illustrate the diversity of approach that I have already pointed to. The coverage is far from complete, and rather than being a handbook we have a collection of contributions by people, mostly from the north-eastern seaboard of the United States, who happen to be doing interesting work.

The most "neuro" of the chapters concern perceptual mechanisms particularly the contribution on perceptual (perhaps "sensory" would be a better word) development by Berkley. In contrast, the excellent overview of cognitive development by Carey is resolutely psychological in tone. Some of the other chapters display what might be called a pseudo-interaction between disciplines. Gazzaniga and Smylie capitalize on the fact that 5 out of the 50 or so split-brain patients have some language function in their right hemispheres. Gazzaniga and Smylie's studies suggest that the presence of language helps the right hemisphere to perform cognitive tasks. However, tasks involving an abstract treatment of language materials are still out of reach of the right hemisphere, being performed, the argument goes, by other cognitive systems in the left hemisphere. Studies of this kind, as with the work on aphasia reported by Zurif, do put constraints on psychological models of normal functioning but this is only cross-disciplinary research in the loosest sense. That we happen to be dealing with the two hemispheres is really incidental; these are psychological studies carried out inside a

theoretical framework which is also psychological.

The same is also true for the work on brain potentials reviewed by Kutas and Hillyard. The existence or not of certain components of these waves in particular tasks can be used to decide between competing psychological theories. However, what it is that is being tapped can only be expressed in terms of the psychological theory that the experimenter adopts. Furthermore it does not actually matter what neural events these electrical components correspond to. The technique is no more physiological than a key press is anatomical.

In general the book is stimulating reading, with chapters ranging through pedagogy in chimps (Premack) and the psychology of reasoning (Johnson-Laird), to a discussion of the mind-brain relationship (MacKay) and an attack on the computer metaphor for mind (Carello and others). It provides a good opportunity to catch up on work and thinking in a variety of areas; but do not expect to find a grand synthesis. □

John Morton is Director of the Medical Research Council's Cognitive Development Unit, University College London.

Winning at physics

Roy Porter

The History of Physics.

By Isaac Asimov.

Walker, New York: 1984. Pp.762.

\$29.95.

TWENTY years ago you might have given your lucky teenage nephew three volumes of Isaac Asimov's *History of Physics* for his birthday. Today you can give it to his son (and nowadays his daughter too) all in one single volume.

Asimov has updated his classic text somewhat to digest recent advances in the subatomic departments, but all the old qualities of the dean of popularization still shine through, bright as a new pin. Above all, Asimov has the great knack of communicating the thrill of science's inexhaustible intellectual energies. He will convince the reader how, once Newton had pondered long and hard about an apple falling from a tree, his mind would thence be "never at rest", till the law of universal gravitation had been hammered out. Story after story unfolds this way of how, pressurized both by logic and by brute experience, the domains of mechanics, light, heat, sound, electricity, and the micro-world of the particles, successively took shape.

Asimov calls his book a history. It is only fitfully that, except in the sense that it shows how certain concepts followed consequentially from others. For Asimov actually takes no interest in the lives of

scientists, in the psychology of discovery and creativity, or in the chronology of progress, let alone in the finer points of historical interpretation. Why was it during the Renaissance that Copernicus advanced the heliocentric hypothesis? — you will look in vain for "Renaissance" here.

In some ways, Asimov's rather Pickwickian notion of history is a missed opportunity. Your grand-nephew or niece won't gather from him what it was that made extraordinary figures such as Kepler and Galileo tick (give them Koestler instead!). And they will get a lopsided view of how science works, one which assumes the only sages worth mentioning are the giants who ended up on the winning side. Boyle, Hooke, Newton etc. all pop up, because of their laws, but even towering intellects such as Francis Bacon, Descartes and Leibniz, who figured no less in the making of scientific thinking and method, are wholly omitted or just mentioned in passing. Asimov is exhilarated by ideas, but they have a curiously disembodied feel to them.

Asimov says he is attracted to the historical approach because, while science changes and grows obsolete, history always stays the same. Well might the chagrined historian riposte: *Eppur si muove*. Yet neither can there be any doubt about Asimov's own mind being in constant motion. For his powers of unfolding the structures of physical thought in a style vivid and familiar, he still has few peers. □

Roy Porter is at the Wellcome Institute for the History of Medicine, London.

Work for revision in geomorphology

Peter Worsley

Geomorphology of Europe.
Edited by Clifford Embleton.
Macmillan Press, London: 1984.
Pp.465. Hbk £45; pbk £17.50.

THE publication of *Geomorphology of Europe* raised hopes that at last there would be a European equivalent to W.D. Thornbury's *Regional Geomorphology of the United States*, or perhaps C.B. Hunt's splendid *Natural Regions of the United States and Canada*. In the event the overall impression is one of disappointment, of a work which has failed to fulfil its potential.

The book was originally conceived as a text to complement the *International Geomorphological Map of Europe*, a project undertaken by the International Geographical Union. The map is as yet incomplete and the book became "a work in its own right", able to stand independently of the forthcoming map. Any critique is naturally deflated to some extent when it is declared that "the editor is well aware of shortcomings in the work and of unevenness of treatment". Indeed one can fully sympathize with the demanding task of managing contributions by some 24 different authors from about ten nations. Apart from the occasional lapse into unfamiliar terminology, Professor Embleton and his Czech colleague Jaroslav Dernek have done wonders in producing a linguistically well-written text.

However one cannot altogether excuse the admitted variability in detail simply because the level of knowledge from region to region is uneven. At the level of generalization appropriate in a work dealing with an entire continent, it ought to be possible to achieve greater uniformity. Further, that "no attempt has been made to standardize the contributions from different authors" is surely a recipe for trouble.

All of the chapters are organized on the basis of structural subdivisions of the continental areas, except for two which deal with the now-submerged shelf. Such an approach is sound but does entail some unfamiliar division of the material; for example the British Isles are discussed in three separate chapters (Caledonian Highlands, West and Central European Highlands, and Hercynian Europe). Within each structural unit summaries are given of the geology, Cenozoic landscape evolution, Quaternary history and coastal character (where relevant), sometimes with comments on contemporary processes. Clearly the newer, climatically influenced geomorphological features do not fit readily into a structural framework but we must recognize the dilemma created by the

need to deal with the entire scope of geomorphology. The greatest benefit from the book will be derived by those who require a briefing on an unfamiliar part of Europe. This would have been aided by the inclusion of key bibliographic citations at the end of each chapter, enabling the reader to quickly identify the main sources available. For instance one is astonished to find no mention in the treatment of England of C.A.M. King's *Northern England* nor A. Straw and K.M. Clayton's *Eastern and Central England*, both of which are extremely valuable guides to the literature.

A close knowledge of each of the various regions is of course necessary to identify inaccuracies. With reference to Britain, we find a strict adherence to a Quaternary framework established over a decade ago, with no hint that this sequence is incomplete or indeed that additional stages have now been defined. The generally jaded flavour is also apparent, for example, in the apparent oblivion to the substantial evidence assembled by J. Boardman for a pre-Devensian glaciation of the Lake District or to J.B. Sissons's penetrating work on the Loch Lomond Stadial cirque glaciation in the same area. One is curious to know where the claimed 300-m-thick crag sequence is located in East Anglia, or where Pliocene crags are subject to rapid coastal erosion, or what is the evidence for a Beestonian sea level of -100 m. In the south-west of England consideration of the important Tertiary Petrockstow Basin is omitted, as it is of the allied major wrench fault which crosses the entire peninsula.

Even less excusable is the poor quality of the illustrations. Many of the maps are reproduced at an inappropriate size (usually too large), the lettering is frequently crude and generally they appear amateurish. Monochrome photographs are dispersed throughout the text, but while some are successful many are badly reproduced. Printing problems apart, the selection of the photographs appears to have been haphazard, many being of dubious value. Perhaps the prize in this respect goes to Fig. 9.36, "Planation surface, probably etchplain with peat cover", which is devoid of any scale and impossible to interpret. The statutory (these days) satellite images are restricted to France and Central Europe and their definition is so poor they serve no useful purpose.

Regrettably, therefore, it is difficult to avoid the verdict that this is a work in need of revision. Such a conclusion is frustrating since the overall objectives are admirable and there is a real need for a synthesis of European geomorphology. This book is a step in the right direction but is, alas, an unsatisfactory first attempt at filling the gap. □

Peter Worsley is Professor of Physical Geography at the University of Nottingham.

Gatekeepers of the chemical cycles

J. M. Hayes

Microbial Geochemistry.
Edited by W.E. Krumbein.
Blackwell Scientific: 1984. Pp. 330.
£35, \$59.95.

THE ocean, about as old as the planet itself and the place of origin of nearly all sedimentary rocks, is one of the main reaction vessels for the processes of global chemistry. Some of the most important actors on that grand stage are microscopic in size. Though each individual works on a scale absurdly small in proportion to the environment, unicellular organisms effectively control the marine chemistry of many elements. Further, bacteria living in the mud and ooze of the ocean floor commonly have the last word before constituents of sediments are more or less immobilized by the process of lithification.

There are fascinating — and still largely uninvestigated, let alone answered — questions regarding the possible importance of bacterial processes in sediments at depths of kilometres and over times familiar to geologists but not microbiologists. In any event, microorganisms again play important roles as the rock cycle is closed by processes including uplift and

**Bring Your
Library Up-To-Date
With This ANS Edition...**

J.M. Cleveland
MONOGRAPH

The Chemistry Of Plutonium

Approx. 680 pages 6" x 9"
Hardbound \$49.00

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for _____ book(s).
☐ Payment in full enclosed.
☐ Bill me and I accept postage and handling charges.

Name _____

Organization _____

Street _____

City _____

State _____ Zip _____

Circle No.01 on Reader Service Card.

erosion. Microbial metabolic activities and chemical by-products accelerate weathering processes and soil formation.

By acting, then, as catalysts and geochemical gatekeepers, microorganisms do much to control the global chemical cycles that shape the sedimentary record. This form of microbial domination persists even today, on a planet teeming with multicellular plants and animals. For most of Earth's history, however, there was no contest; all biology was microbial.

For those good reasons, geochemists concerned with low-temperature processes ought to be as interested in bacteria as in the source of their next pay cheque. Some good books are now appearing in the prefix-laden field of biogeochemistry, but the subject draws upon so many disciplines that few authors are attempting large-scale syntheses. In 1979 T. Fenchel and T.H. Blackburn provided us with a superb textbook (*Bacteria and Mineral Cycling*, published by Academic Press), and now Wolfgang Krumbein, a most engaging exponent of modern studies in geomicrobiology, has assembled an outstanding group of authors to review aspects of microbial geochemistry.

The volume includes some particularly notable contributions. K. H. Nealson provides beautifully documented reviews of the microbial iron and manganese cycles, J. Berthelin discusses microbial weathering processes in similar detail, and a chapter by C. D. Curtis outlines the roles of microorganisms in diagenesis of sediments. In each case, diverse lines of evidence have been pulled together to provide a unique and authoritative view of the subject.

The Institute of Ecology and Genetics at the University of Aarhus, Denmark, is currently one of the most active centres for research in sedimentary microbiology. To the large number of excellent papers coming from that laboratory we must now add chapters in this book by T. H. Blackburn, on microbial nitrogen cycling, and by B. B. Jorgensen on microbial sulphur cycling. The former emphasizes mineralization — a welcome change from most other reviews — and the latter helpfully expands on previous treatments, although, falling as it does in a particularly active area, it suffers slightly from the evident lapse of time between its preparation and the appearance of the book.

An otherwise lucid review of microbial palaeontology written by A.H. Knoll and S. M. Awramik suffers even more strongly from the same problem. The volume is completed by a discussion of microbial carbon cycling by Krumbein and P. K. Swart, and a very nicely illustrated chapter on silica cycling, by Krumbein and D. Werner. □

J. M. Hayes is Professor of Biogeochemistry in the Departments of Chemistry and Geology at Indiana University, Bloomington.

Renewal of surface activity

Pasupati Mukerjee

Surfactant Systems: Their Chemistry, Pharmacy and Biology.

By D. Attwood and A.T. Florence.

Chapman & Hall/Methuen: 1983.

Pp. 794. £45, \$99.

SURFACTANT science, in its widest sense, is a part of a large number of academic disciplines, and deals with a wide variety of biological and other natural systems, and industrial operations. The original objective of the authors in writing this book was to revise a useful monograph, *Solubilization by Surface-Active Agents* by P.H. Elworthy, A.T. Florence and C.B. McFarlane, published in 1968 but now out of print. This revised version, however, is a much larger volume and of considerably greater scope. Surfactants are defined rather broadly — typical amphipathic substances with flexible chains such as soaps and detergents, and membrane lipids are included, as also are physiological surfactants, such as bile salts, and a variety of drugs which display interesting surface-chemical and colloidal properties.

The authors' coverage of the various sub-topics is relatively unorthodox. The first three chapters consider general aspects of surface activity, phase behaviour of surfactants and problems of self-association and micelle formation. Chapter 4, the longest in the book, deals with the surface activity and colloidal properties of drugs and some naturally occurring substances. Solubilization is discussed as a general phenomenon and its areas of pharmaceutical application are covered. Chapter 7 is devoted to some implications of the role of surfactants in formulations, including their effects on

drug dissolution, membrane permeability and drug absorption, after which there are two chapters on emulsions and suspensions and an account of some selected toxicological aspects of surfactants. The concluding chapter, on reactivity in surfactant systems, includes a brief treatment of aspects of chemical and photochemical reactivities of solubilized species as well as of the chemical stability of surfactants themselves. Much emphasis has been placed on commercial non-ionic surfactants, which are chemically heterogeneous but which are used extensively because of their compatibility with physiological systems.

The book is aimed at final-year students of pharmacy and at

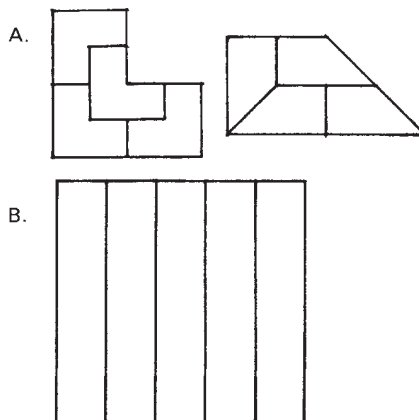
... post-graduate students of pharmacy, biochemistry, biology, chemistry, and those working in industrial research and development laboratories exploring the value or problems of surfactant systems.

Yet I do not believe this volume can be described either as a good textbook or as a critical monograph. The literature is covered in too descriptive a manner — in many cases where an evaluation of opposing points of view might have been desirable, the authors have merely stated those points of view and have left readers the task of making their own conclusions. On the other hand, the literature has been covered extensively. There are 2,427 references at the end of the chapters, only a small fraction of which are repeated. So while I feel that the authors have fallen short of their primary goal, the book is likely to be of considerable value for reference and as a review of the literature, particularly for those who are interested in surface activity of drugs and the formulation and use of medicinal products containing surfactants. □

Pasupati Mukerjee is a Professor in the School of Pharmacy at the University of Wisconsin-Madison.

Divisive difficulties from David Singmaster — the solutions

Answers to the three problems appearing on p. 521 of Nature, 9 August.



The preceding question was designed to mislead you! It did *not* say the areas must be congruent to the original area.

Here the square is divided into five areas, but each area has two parts. The question did *not* say that the areas must be connected (but did state that many people would find this solution unacceptable!). If one insists on connected areas, it is not known whether there is any way to divide a square into five (or even three) congruent parts.

For the chemistry laboratory

Twenty-five new products from American companies, with the emphasis on speed and precision.

● The Model 220 solvent delivery system from Scientific Systems Inc. is designed to perform equally well as a dependable system for routine analyses, or as a research instrument that can be upgraded to gradient operation. The Model 220's flow rate can be set in 0.01 ml increments from 0.05 to 9.99 ml per min. The compressibility compensation control allows adjustment when extremely accurate flow rates are required, and the wide dynamic range pulse damper keeps pulsations to a minimum. The prime/purge valve and Luer-type drain port, both conveniently located on the front panel of the Model 220, make priming and purging of the pump a simple operation. Digital upper and lower pressure limits automatically stop the pump when beyond the set range.

Circle No. 100 on Reader Service Card.

● A new HPLC absorption monitor with very high sensitivity (0.001 A for full-scale recorder deflection) is now available from Knauer. Its high stability is accomplished by using a low pressure mercury cathode and the unit features a microprocessor-controlled auto zero for the entire sensitivity range. The detector also features a bright digital readout of absorbance for ease of operation. A wide choice of flow cells is available ranging in size for use in micro-bore through to preparative chromatography. An accessory lamp is available for operation in the visible region of the spectrum.

Circle No. 101 on Reader Service Card.

● The Gilson 401 dilutor is a microprocessor-controlled single syringe instrument which is user programmable in an easy conversational mode to store up to 99 programmes with battery protection in the event of a power failure. It uses only four syringes to cover the range 0.2 µl to 10 ml with a maximum operating speed of 1.5 s using a stepper motor drive for high precision and reproducibility, even with small piston displacements. Chemical inertness is assured by using a glass syringe with PVDF piston and ceramic valve. The five function keys directly address the basic programmes; dilution, dispensing, pipetting, repartition, and link each with a capacity of up to 12 samples per programme. The 401 has been designed for use with automatic programmable liquid handling procedures in clinical laboratories such as RIA, EIA, enzymology, microbiology and serology. In addition it can be used as a slave to an external computer or other Gilson liquid handling products such as the new sample preparation unit and sample changers in industrial and research laboratories.

Circle No. 102 on Reader Service Card.

● The multiple casting chamber, SE 215, is a new accessory to the Hoefer miniature slab gel unit, the Mighty Small SE 200. The SE 215 casts up to ten 1.5 mm thick gels simultaneously. Multiple casting of gels reduces preparation time and ensures uniformity in pore size and/or gradient shape. The chamber is made of sturdy acrylic and the front panel is removable to facilitate assembly and disassembly of the stack of gel sandwiches. Gradient gels are poured from the bottom. The unit comes with ten notched alumina plates, twenty 8 × 10 cm glass plates, and ten sets of T-shaped spacers in each of three thicknesses: 0.5 mm, 0.75 mm, and 1.5 mm. Combs are available for use with the SE 215.

Circle No. 103 on Reader Service Card.

● The new "Economy" tray dryer from Labconco is a lyophilizer which performs bulk or batch drying. The drying chamber is made of durable fibreglass and a three-inch vapour outlet at the base of the chamber allows connection to a Labconco freeze dryer-18, -8 or -5. The chamber is equipped with three 12 × 12 inch interior shelves. Each shelf has a heater embedded in the underside with temperature probe and all shelves are adjustable to accommodate various sizes of containers. The dryer is equipped with a separate shelf temperature controller and three probes provide accurate temperature control on each shelf. A vacuum bleed valve mounted on the side of the drying chamber provides controlled release of vacuum.

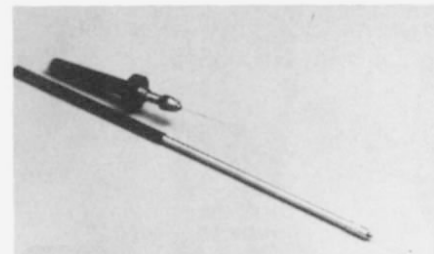
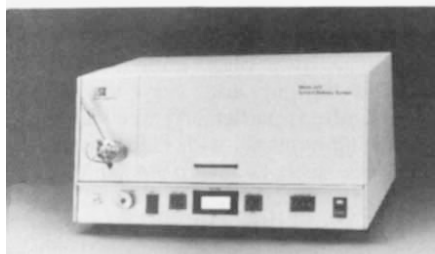
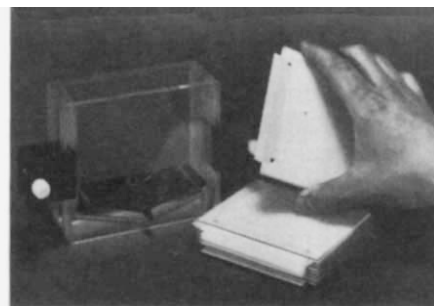
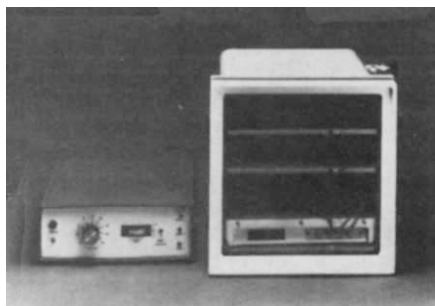
Circle No. 104 on Reader Service Card.

● Chemtix Inc. of Hillsboro, Oregon, has produced an intriguing range of devices called SMORCs, standing for small orifice cleaners. The 4-inch short-handled SMORC features a pin vice to securely hold fine cleaning wires and a hollow handle for storing the wires; it is ideal for cleaning recorder pens, syringe needles, burettes, valves and so on. The 10-inch, long-handled SMORC is designed to fit into hard-to-reach areas for cleaning. Each instrument comes with three sizes of wire and a magnet surrounding the handle so it can be placed on steel machinery or cabinets to avoid loss.

Circle No. 105 on Reader Service Card.

● New from Labconco is a digestion unit which digests up to 25 samples in half the time of conventional methods. The Rapid Digester will digest material such as feeds, grains, plant tissue, water effluent, organic wastes and food products for protein/nitrogen determination. The Rapid Digester has 25 individual heaters constructed of ceramic materials and these are encased in insulating material to prevent heat loss and assure low energy consumption. A solid-state controlling unit regulates temperature, and monitoring lights indicate when the digester reaches its preset temperature. The Rapid Digester comes with a carrier rack for easy handling of digestion tubes and heat shields prevent heat loss when in place, and permit quick cooling when removed. A fume removal system is also available as an accessory to eliminate hazardous fumes.

Circle No. 106 on Reader Service Card.



Labconco tray dryer (top left), Hoefer multiple casting chamber (top right), Scientific Systems solvent equipment (bottom left) and Chemitrix small orifice cleaners (bottom right).



FOXY matches fractions to your chromatogram. Not to the clock.

Program your collection by absorbance peaks, retention times, timed intervals, counted drops, or pumped volumes. You can even enter fraction size directly in ml if you have an ISCO WIZ pump. FOXY rejects void volumes, locates and saves desired peaks, controls valves and pumps to make gradients or automate your system, and changes racks and volumes according to peaks or time. For HPLC, collect successive injections without re-setting, and overlay repeated injections of the same sample to collect identical peaks in the same tubes. The sloped keyboard, 12-digit display, and logical programming make it all easy to do.

To learn more, call toll-free [800] 228-4250. Or write ISCO, P.O. Box 5347, Lincoln, NE 68505.



Circle No.03 on Reader Service Card.

- Designed for use in clinical, biomedical, research, industrial and academic laboratories to supplement the familiar fume hood, Mystaire fume enclosures from Heat Systems-Ultrasonics return purified air to the room. This saves on heating and cooling expenditure, and, since no ducting is required, also on construction and rebuilding costs. The system includes a primary mechanical filter stage with electrostatically charged fibres to trap submicrometre particulates at an efficiency of 99.95%. The main filter is an activated charcoal or chemisorptive filter, or specially treated filters for radioactive iodine or mercury. A final polishing stage consists of carbon-impregnated cloth, and a brushless, sparkless motor drives the corrosion-resistant fan; all electrical connections are isolated from the airstream for safety. Two sizes are available and the enclosures may be car mounted for mobility.

Circle No.107 on Reader Service Card.

- The new Spectra-scope from Barnes Analytical/Spectra-Tech is a useful in-compartment instrument that speeds and simplifies sampling and viewing with FTIR spectrophotometers. With the Spectra-scope, samples of any substance from micro range to macro sizes can be mounted and viewed immediately in FTIR spectrophotometers, with no need for external detectors; samples do not require cutting or special preparation, in contrast to conventional microscopy techniques. Although only about 7 inches tall, the new sampler includes many of the high-performance features of IR microscopes: image masking, micro positioning, interchangeable reflecting objectives, flexible illumination, laser block and purge.

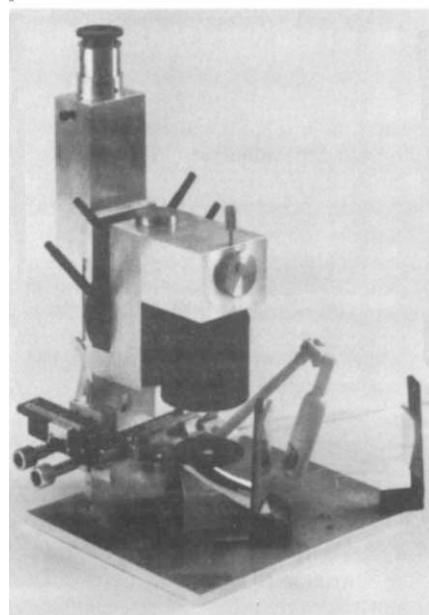
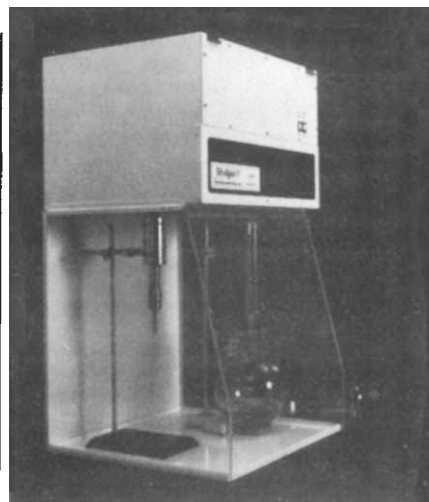
Circle No.108 on Reader Service Card.

- Two new models have been added to the Allen Datagraph 600 Series of portable, battery powered x/y recorders. The Model 625 offers a built in x axis time base with eight ranges that allow YT recording. The Model 635 has the same time base plus calibrated offset controls for both axes. A 1,000 count dial provides voltage offsets up to 100 scales depending on the range switch setting. This permits very small signals riding on large voltages to be recorded. All 600 Series models feature English or metric calibration, battery or mains operation, rechargeable gell cell battery and a hysteresis charging circuit. Twelve fixed recording ranges are provided with both modes.

Circle No.109 on Reader Service Card.

- The World Precision Instruments model FD223 dual/differential electrometer, with high impedance, low leakage current and very low input capacitance, is particularly well suited for measuring potential with high source impedances, such as macro and micro ion selective and pH electrodes. Two miniature gold-plated probes with driven guard to minimize input capacitance and leakage current are included.

Circle No.110 on Reader Service Card.



Mystaire fume enclosure (top) and Spectra-scope from Barnes Analytical.

- PolyScience Corporation Chemical Division has produced a new analytical standards kit, containing fourteen naphthenes — saturated cyclic paraffins in the boiling range 30° to 200°C. The substituted cycloparaffins exist as *cis* and *trans* isomers which increases the number that can exist compared to substituted straight-chain paraffins. These compounds have the same boiling range as "full range naphthas" that are used in making gasoline. This kit, No.281C, together with kits 211C (*n*-paraffins), 221C (branched paraffins), and 251C (benzenoid), contain many of the compounds found in these naphthas. The new 281C calibration mixture contains equal parts (by weight) of *n*-pentane (C₅) through *n*-dodecane (C₁₂) and are of the same boiling range as the kit compounds. With this mixture it is possible to assign Kovats numbers to be used in identification of unknown peaks. The liquids (2ml each) are sealed in glass ampules for shipping; labelled screw-top bottles are included for storing them after the ampules have been opened.

Circle No.111 on Reader Service Card.

● Dapple Systems, Inc. has introduced the Autoscan scanning densitometry system for quantitative Autoradiography and the analysis of two-dimension films, filters and gels. A high resolution video camera coupled directly to a light microscope or simply viewing a photograph produces the video image which is then digitized into 64 grey levels for analysis. Automatic measurement of spot location and brightness is accomplished by setting discriminators to separate the spots from background. Alternatively, irregular-shaped areas may be traced directly on the video monitor and these total areas scanned for brightness. Density is determined by comparing the measured brightness to a stored calibration curve. Statistical analysis of the area, perimeter and density of the outlined features is also possible. The system includes a 128 Kbyte Apple IIe computer, dual floppy disk drives, monitor, dot matrix printer, video camera and cursor control touch pad.

Circle No. 112 on Reader Service Card.

● With the introduction of the new AccuSpin and Accuspin FR centrifuges, Beckman Instruments, Inc. are bringing the latest technology to tabletop centrifugation. Digital controls and displays provide accuracy and show run conditions at a glance. Diagnostic messages appear if something needs attention. Programmed acceleration provides optimum acceleration speeds for fast sample processing, while offering a gentle start to protect gradients. A digital integrator allows precise speed control, accurate to 10 r.p.m., with no overshoot. The refrigerated model, the Accuspin FR, has an advanced frost-free refrigeration system, which eliminates ice and water in the chamber. Safety features include self-sealing rotor buckets, automatic door interlock, rotor imbalance detector and aerosol containment covers; these covers allow users to spin almost every size of laboratory tube with reliable protection from aerosols. Three rotors are available for the digital AccuSpin tabletop centrifuge: a horizontal rotor with four aluminium swinging buckets and two fixed angle rotors. The horizontal AH-4 rotor reaches speeds and forces to 4,200 r.p.m. and 3,200g; the fixed angle rotors reach 4,900 r.p.m. and generate 4,800g-force.

Circle No. 113 on Reader Service Card.

● Vector Laboratories is now offering a new fluorescein labelled lectin kit, providing a panel of seven fluorescent lectins with a range of carbohydrate specificities. This kit consists of 1 mg of each of the following lectins: *Bandeiraea simplicifolia* (BSL I), *Lens culinaris* (LCA), *Phaseolus vulgaris* leucoagglutinin (PHA-L) and erythroagglutinin (PHA-E), *Pisum sativum* (PSA), *Sophora japonica* (SJA) and succinylated wheat germ (S-WGA). Each lectin is labelled with an optimal number of fluorochromes to provide maximum staining.

Circle No. 114 on Reader Service Card.



Dapple densitometry system (top left), new filter from Sartorius (top right), and Micro-Plan II image analysis system from Laboratory Computer Systems.

● The Jouan Inc. C3000 ventilated and the CR300 refrigerated tabletop centrifuges combine high speed operation up to 6,000 r.p.m. and RCF of 5,840g with an 1.8 litre capacity to process more samples per run. The high g force significantly reduces run time necessary to achieve separation. The centrifuges are available with either horizontal or fixed angle rotors and a full range of inserts, allowing centrifugation of tubes and bottles ranging from 250 microlitre to 450 millilitre. A special carrier is available for centrifugation of microtitre plates. Standard features include programmed acceleration and continuously variable dynamic braking which allows the centrifuge to start gently, achieve speed quickly, and allows the operator to select braking rate thus preventing re-suspension of samples in even the most sensitive gradients. Safety features include an impact-resistant stainless steel guard bowl with no protruding sensors, a steel guard ring and a unique lid lock mechanism which prevents any operator from opening the lid while the centrifuge is in use. The imbalance detection system automatically shuts down the unit in an imbalance situation. The refrigerated model has temperature control from 5°C to 25°C and uses a ¼ h.p. true refrigeration system. Speed and temperature readouts allow operator to monitor actual values at all times.

Circle No. 115 on Reader Service Card.

● For small volume syringe clarification of both aqueous solutions and aggressive solvents, the Minisart-S filter from Sartorius offers convenience and reliability for HPLC and other work. The new filter is designed for minimum dead volume, high flow rates and absolute 0.45 micrometre retention.

Circle No. 116 on Reader Service Card.

● An addition to its Micro-Plan II image analysis system has been announced by Laboratory Computer Systems Inc. Statistical analysis software is now included with the Micro-Plan II for use with the IBM PC or PC/XT. The Micro-Plan II is a low cost, tablet-oriented image analyser that performs measurements such as length, area, centre of area, form factor, feret x-y, and maximum diameter of two dimensional features traced on the digitizing tablet. The software allows the Micro-Plan II user to analyse these measurements statistically and to create data files in special formats that can be used for publication or read by Lotus Development Corporation's spreadsheet program 1-2-3. The statistical analysis software will accept data directly from the keyboard, as well as from the Micro-Plan II, to allow correlation studies between morphometric measurements and other experimental variables. The software is provided as public domain source code so that it can be modified by researchers. The Micro-Plan II, in conjunction with an IBM PC or PC/XT, provides a flexible solution to quantitative image analysis problems.

Circle No. 117 on Reader Service Card.

● Kirby Lester announces the new Microlink modular interface for connecting laboratory and monitoring equipment to microcomputers. The system operates with IBM, Hewlett Packard, Apple, Commodore, Victor and others. Microlink consists of a mainframe with a power supply and IEEE 488 interface. Up to 18 modules can be plugged into the system and there are currently 30 modules to choose from with 6 more planned this year. Included in the range are analog signal conditioning, high speed data acquisition, timers, and analog to digital converters.

Circle No. 118 on Reader Service Card.

● The Jewett range of refrigerators provides safeguards for explosion proof operations in critical applications. The models range from the more economical, which provide explosion safe capabilities, to sophisticated units offering total explosion proof construction for any situation. These refrigerators are suitable for petrochemicals research, or in any situation where highly volatile substances are handled. Thin-wall design of the explosion-safe LR series models LR12C, 17C, 37C, and 55C, increases storage capacity, while interior surfaces of 20-gauge stainless steel ensures explosion safe conditions due to non-sparking surfaces and fixtures. These refrigerators are available as custom-finished, upright models with single and double doors. Undercounter and wall mounted units with similar features are also available. Three totally explosion-proof refrigerators, models UL5CE, LR12CE and RE200E, are also available, prohibiting combustion whether air inside or outside cabinets contains flammable vapours. Designed for use in hazardous locations, they feature exterior-mounted fully ventilated condensers, and automatic condensate evaporators which eliminate the need for floor drains. All controls and wiring are external and factory-sealed in approved conduits and custom cast aluminium housing, providing complete explosion proof protection from sparking or arcing in electrical wiring relays, thermostat switches or compressors. They are available in undercounter versions, 4.9 cubic feet, and single and double upright models, 11.8 and 21 cubic feet.

Circle No. 119 on Reader Service Card.

● Boehringer Mannheim Biochemicals has expanded its endoproteinase line to include Glu-C (*Staphylococcus aureus* V8 protease) and its specific chromogenic substrate, carbobenzoxy-phenylalanyl-L-leucyl-L- α -glutamyl-4-nitranilide (CPLGN). This serine protease has become a valuable tool in protein sequencing. In ammonium bicarbonate buffer (pH 7.8), or ammonium acetate buffer (pH 4.0) endoproteinase Glu-C will specifically hydrolyse peptide and ester bonds at the carboxyl side of glutamic acid. However, in the presence of phosphate buffer at pH 7.8, this endoproteinase will cleave at the carboxyl side of both aspartic acid and glutamic acid.

Circle No. 120 on Reader Service Card.

ADVERTISEMENT

**PHORBOL
ESTERS**

**PHORBOL ESTER
RESEARCH KIT**

PMA PDD POBZ PDA
 MPMA 4xPDD PHORBOL
 Cat. #P1327 \$165.00

LC SERVICES CORPORATION
165 New Boston Street
 Woburn, MA 01801
 (617) 938-1700

Circle No. 22 on Reader Service Card.

● A new 1700 Series of Fourier transform infrared spectrometers has been introduced by Perkin-Elmer. Based on a single high performance interferometer, the instruments in the 1700 Series have special data handling capabilities. The spectrometer itself uses an improved Michelson interferometer designed to compensate internally for any effects of dynamic misalignment. The result is a highly stable system which is resistant to vibration and disturbances during a scan. Interferometer alignment is very simply carried out by a closed-loop servo system and is initiated from the control console. The standard spectrometer has a frequency range of 4,400 cm^{-1} to 400 cm^{-1} and a choice of resolution down to 2.0 cm^{-1} . The resolution can be increased to 1.0 cm^{-1} by the addition of an optional memory board. The noise level is good, making the 1700 Series ideal for difficult, energy-restricting samples. At around 2,000 cm^{-1} at 4 cm^{-1} resolution, the peak-to-peak noise level is better than 0.1%T for a 4-s measurement. The standard instrument has a second sample position adjacent to the detector, which can increase the throughput by a factor of 10 for samples that are highly scattering. Stability is a feature of the 1700 Series design. Both the air cooled source and the fast response DTGS detector are actively stabilised and the optical path is sealed and desiccated for high background stability.

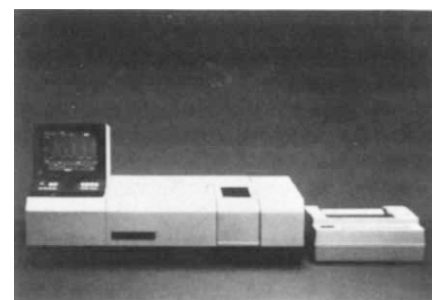
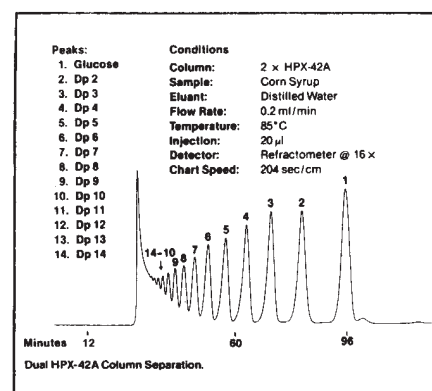
Circle No. 121 on Reader Service Card.

● Bio-Rad's Aminex HPX-42A carbohydrate analysis column is dedicated to fast, high resolution analysis of polysaccharides. Packed with a 4% crosslinked silver form resin, a single 300 $\times$ 7.8 mm column can resolve glucose polymers as large as Dp-11 within 30 minutes. The primary separation mechanism is size exclusion, although ligand exchange affects selectivities for monosaccharides. A single Aminex HPX-42A column provides excellent resolution of carbohydrate polymers up to Dp-11, and for better resolution of polysaccharides greater than Dp-11, two columns may be used in series. The use of two columns allows resolution previously only possible with gel filtration chromatography.

Circle No. 122 on Reader Service Card.

● A new cryoscope, the Wide Range Cryette from Precision Systems, Inc. can determine freezing point of 2 ml samples in most solvents with freezing points in the range of -6 to +22 $^{\circ}\text{C}$. These data can be used to determine molecular weights of up to 1,200, as well as to indicate the concentration of solutions. Freezing point is determined with precision of 0.001 $^{\circ}\text{C}$, which is sufficient for molecular weight determination of around 1% precision. The Wide Range Cryette is also useful for making quick checks of reagent and solution purity, and can save time in monitoring the output of a chromatography column.

Circle No. 123 on Reader Service Card.



A trace from the Bio-Rad Aminex HPX-42A (top) and the Perkin-Elmer 1700 infrared spectrometer.

● MicroCal's new MC-2 differential scanning calorimeter utilizes two thermopile-driven feedback systems for ultra high sensitivity. Designed to operate over the temperature range from -20 to +110 $^{\circ}\text{C}$ with dilute aqueous or non-aqueous solutions, the instrument is ideal for studying structural transitions of proteins, polynucleotides, lipids, and membranes. In addition, an accessory cell is available for solid samples. The usable sensitivity is substantially improved over the earlier MC-1 model; protein transitions often can be studied quantitatively at concentrations below 0.1%, lipid transitions at concentrations below 0.01%, etc. The matched 1.3 ml sample and reference cells are of the total-fill type with access only through external filling ports. Cooling of the cells is accomplished by circulating fluid from an external bath through veins in the adiabatic shield. In conjunction with MicroCal's DA-2 data acquisition and analysis system, the analog instrument may be interfaced to an IBM Personal Computer for automated operation, colour display of data, hard-copy printout, and data storage and analysis. State-of-the-art software also provides the capability for extracting thermodynamic parameters for transition, curve-fitting to Van't Hoff equations, and deconvolution of overlapping peaks.

Circle No. 124 on Reader Service Card.

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.